Building Blocks for 64-bit AMD Opteron ™ Processor-Based Clusters

Rich Brunner
AMD Fellow, SW Architecture

ClusterWorld June 24, 2003

Review: Scale-Up

One system, [mid-range is typically SMP configuration]:

- Multiple processors (up to 16P)
- Large amounts of shared memory
- Shared I/O resources
- High bandwidth and low latency system interconnect
- OS must handle the balancing of incremental resources causing the complexity of OS to scale with processors and memory

- Critical that the system design supports balanced scaling of processors, memory, and I/O devices
- Fits a number of legacy commercial application scenarios where close sharing of memory by multiple worker threads is critical.
- 2P system that incrementally scales to 16P can cost 5x that of a 2P system that scales to 4P.
  - You pay now for the ability to scale-up later.
  - Systems that scale-up beyond 4P are more costly due to the complexity of the system, its interconnect, and RAS features.

Review: Scale-Out

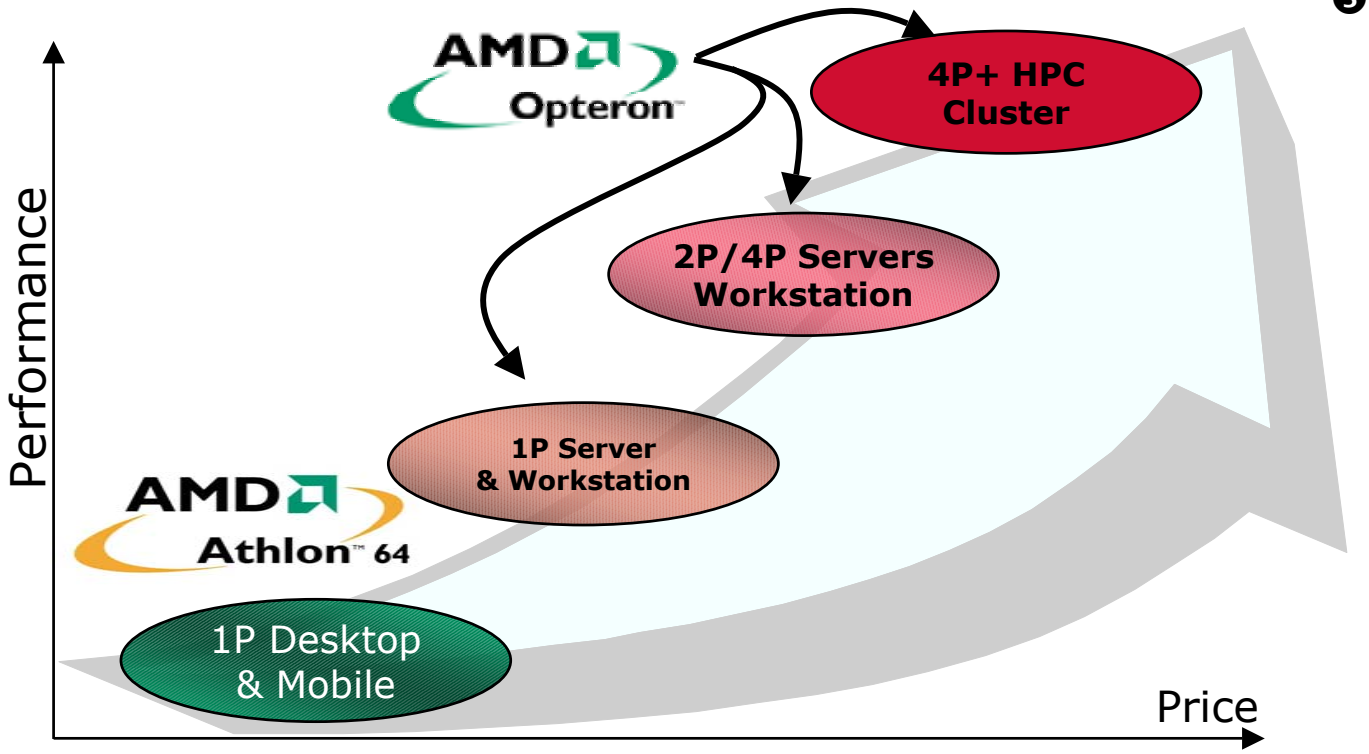
Many small, simple systems:

- Each system has 2 to 4 processors
  - Each system has its own memory
    - Systems are interconnected by a common network

- Each individual system does not need to be expensive, complex, and high-tolerance design
- Small size allows modular and dense packing in racks – very flexible
- Many RAS requirements can be satisfied by simple redundancy and ease of replacing/upgrading individual
- Network parallelism much easier for standard OS to handle
- However, the network is relatively slow interconnect:
  - Good fit for applications/workloads which can be decomposed into multiple threads/tasks with little data-sharing or communication.
  - Fits scientific/technical computing well

Targeting the Benefits of Both

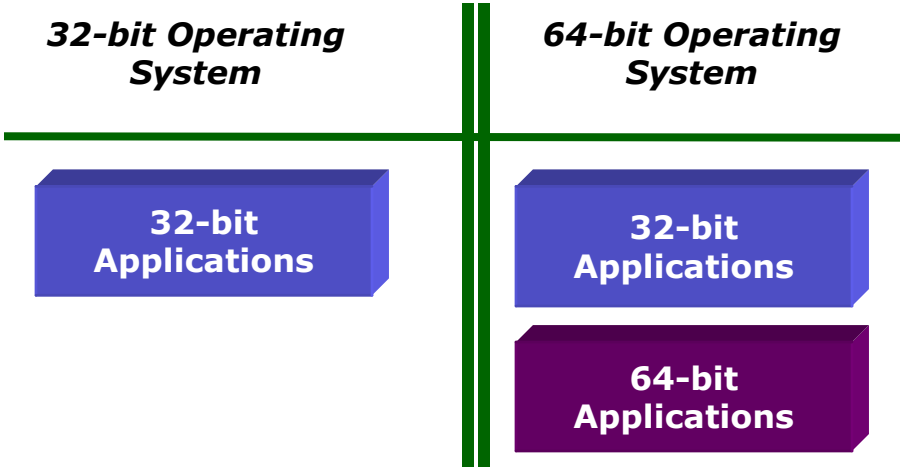
AMD is targeting the mid-range, high-volume server space by supporting the best of both scale-up and scale-out approaches

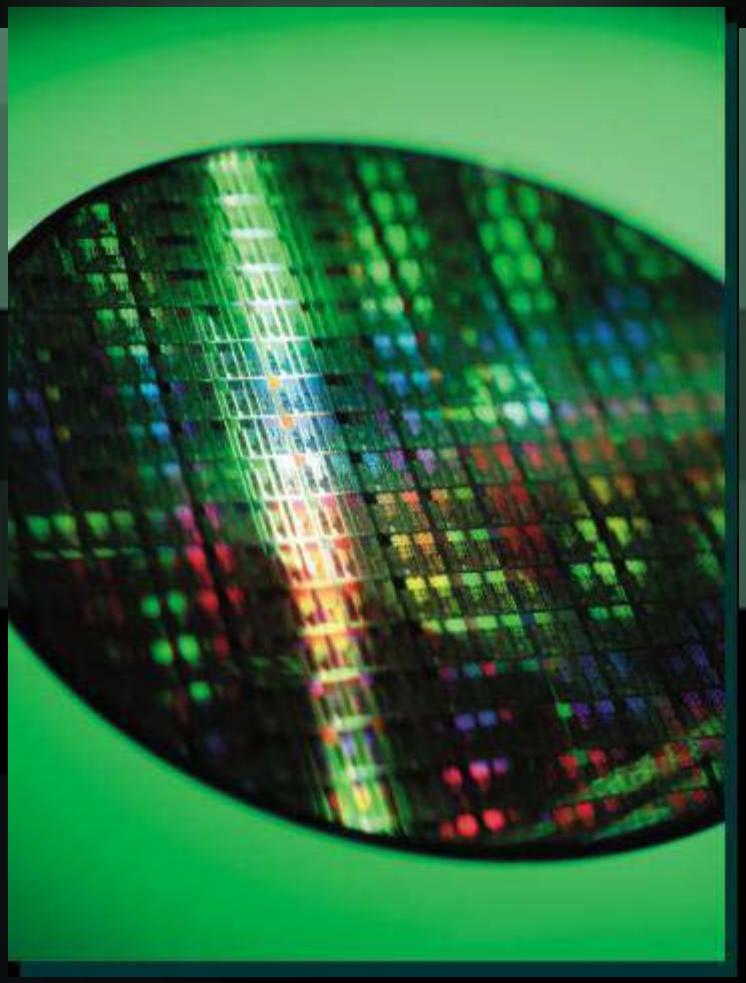


❺ Leverage & extend the Industry investment and expertise in x86

Building a Bridge from the 32- to the 64-bit World

- Leverages the initial success of AMD Athlon™ MP processor
- Adds 64-bit capabilities to the world's highest performing 32-bit core for 2P and 4P servers
- Current 32-bit applications will work on both 32-bit and 64-bit operating systems
- Doesn't require special hardware or investment in a proprietary infrastructure
- Developing a solid ecosystem of motherboards, operating systems, development tools, and device drivers





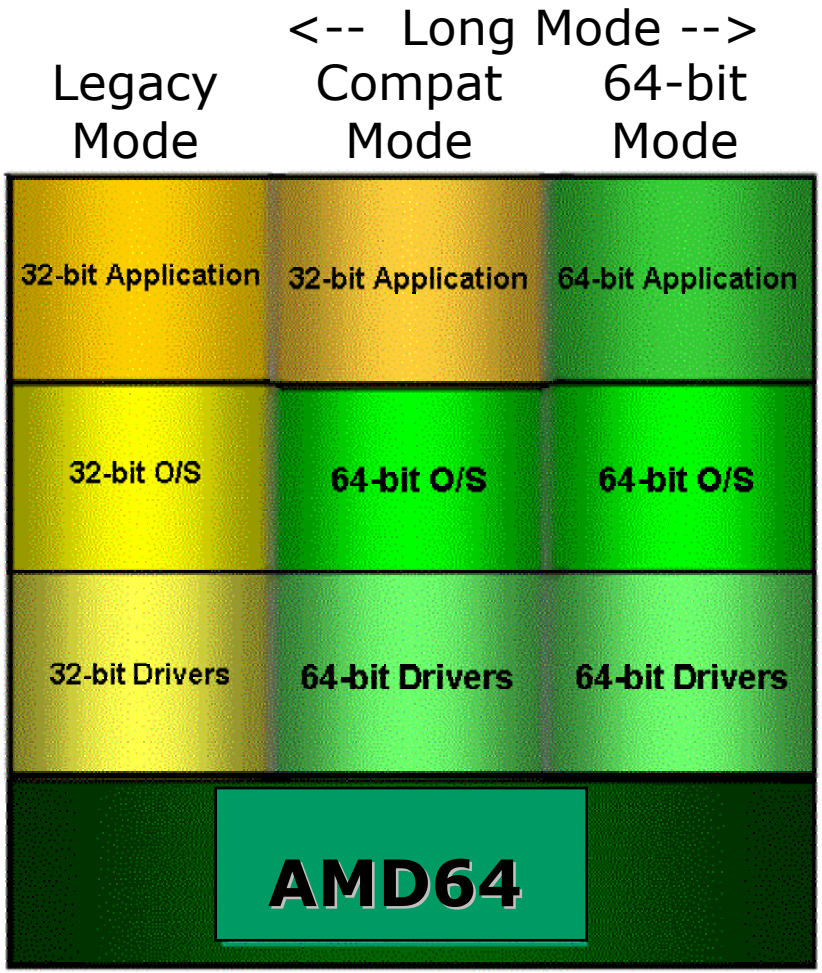
AMD64 Architecture

AMD64 Computing Strategy (1)

AMD took the x86 architecture and extended it to 64-bits to make the AMD64 architecture

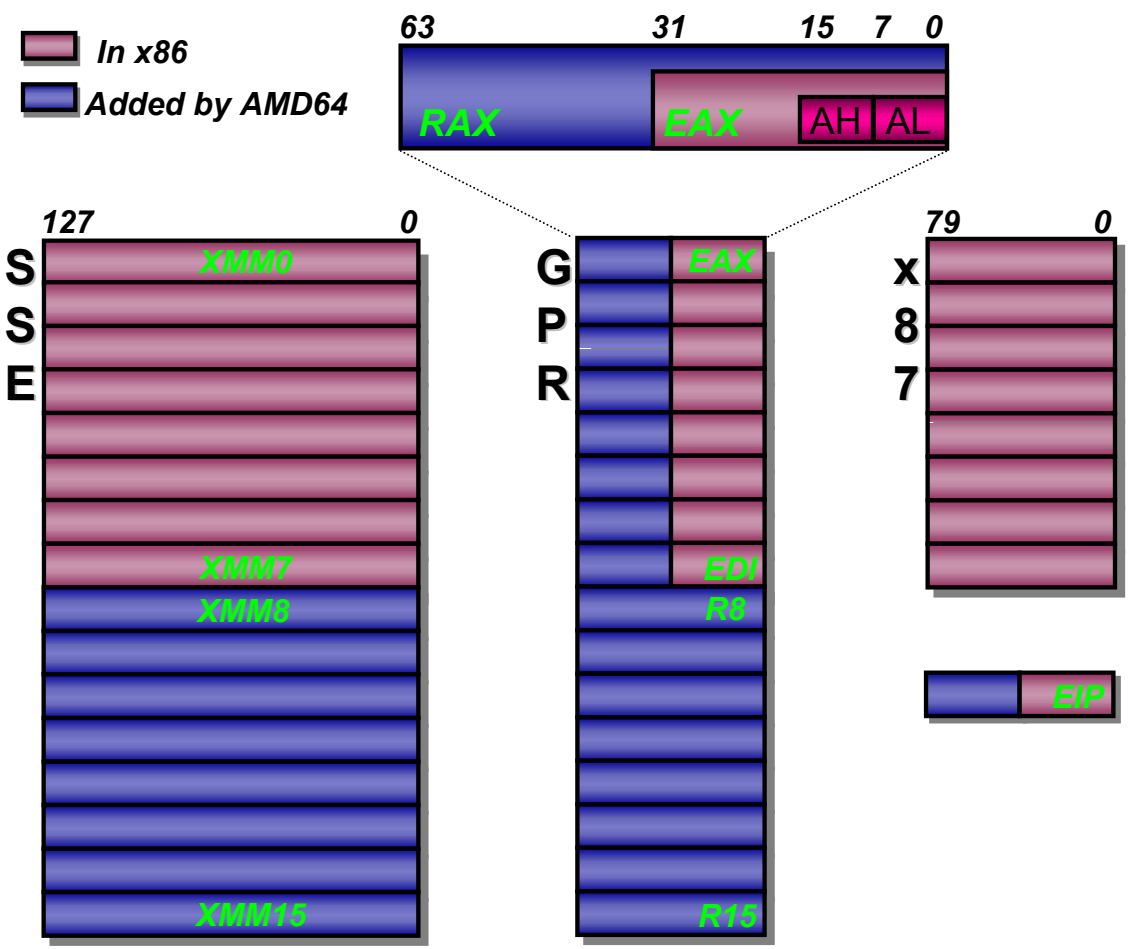
Extensions are so simple and compatible, that the processor can support both x86-32 and x86-64 at full speed & performance

- Offers compatibility and performance for 32-bit applications (legacy mode)
- Can move to 64-bit addressing and data types without giving up 32-bit compatibility (long mode)
  - Leverages the key PC infrastructure rather than needing to re-invent it.



AMD64 Computing Strategy (2)

- AMD64 Architecture:
  - 64-bit integer registers
  - 64-bit Virtual Address
  - 52-bit Physical Address
  - Sixteen 64-bit integer regs
  - Sixteen 128-bit SSE regs
  - SSE2 Instruction Set
  - Double precision scalar and vector operations
  - 16x8, 8x16 way vector packed integer operations
  - SSE1 already added with AMD Athlon™ MP Processor



64-Bit Mode Operation (1)

- Default data size is 32 bits
  - Override to 64 bits using new REX prefix
  - Override to 16 bits using legacy operation size prefix (66h)
- Default address size is 64 bits
  - Pointers are 64 bits
- 2 New instructions added, Some redundant encodings reclaimed
  - MOVSXD: Move sign extended double to quad
  - SWAPGS: Allows quick swap of GS in ISRs
- New override (REX) allows naming 16 GP and 16 SSE registers
  - Only 1 override byte per-instruction is needed for extended registers; regardless of how many are used by the instruction

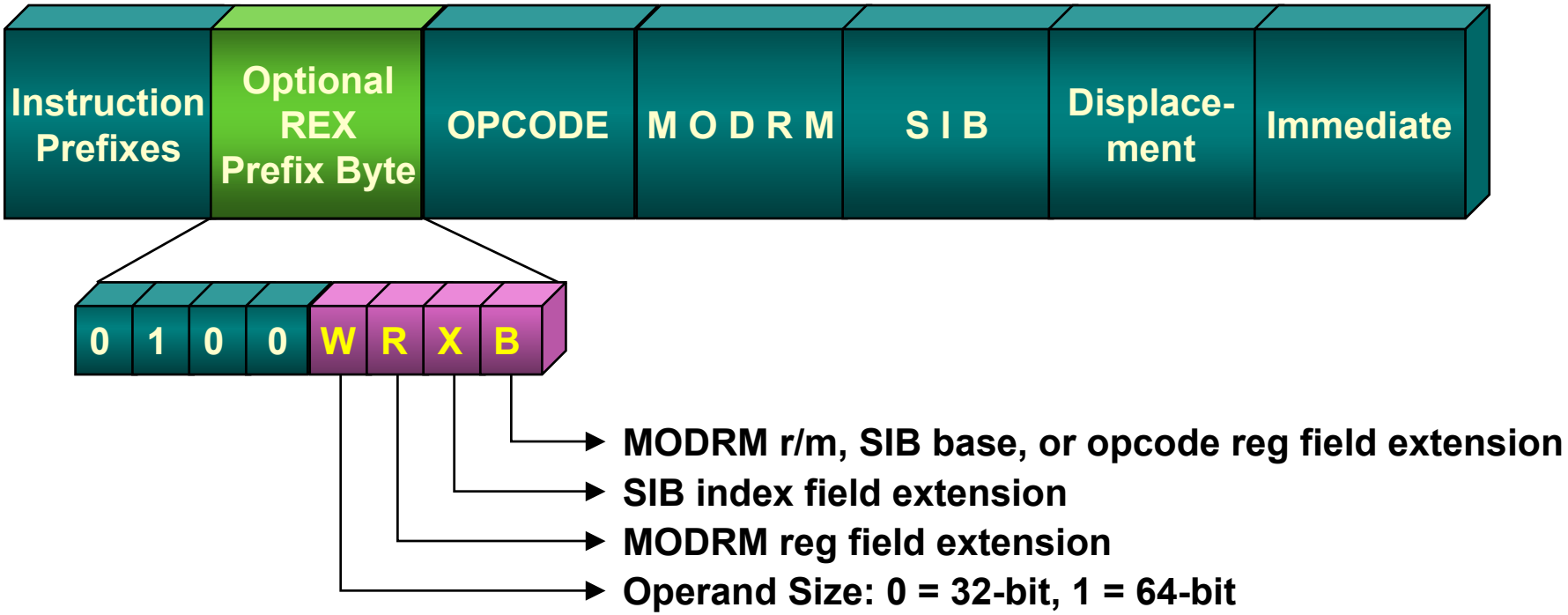
Prefix Type	Default	REX	66h
Operand Size	32	64	16

64-Bit Mode Operation (2)

- Paging extended to 4-levels to provide 64-bit addressing
  - Page Table entries simple extension of x86 PAE-formatted entries
  - AMD64 supports 64-bit Virtual Address & 52-bits Physical Address
  - AMD Opteron™ Processor implements 48-bit Virtual Address and 40-bit Physical Address
- Interrupts and exceptions create 64-bit state
- 64-bit mode uses flat, unsegmented virtual address space
  - The legacy x86 segmentation scheme is disabled in 64-bit mode
- Code Segments still exist in Long Mode to specify default mode (16-, 32- or 64-bit) and execution privilege level (CPL)
  - So existing privilege level and checking mechanisms are retained
- Switch between 64-bit mode & Compatibility Mode accomplished via normal Far Transfer instructions:
  - CALLF, RETF, JMPF, IRET, INT

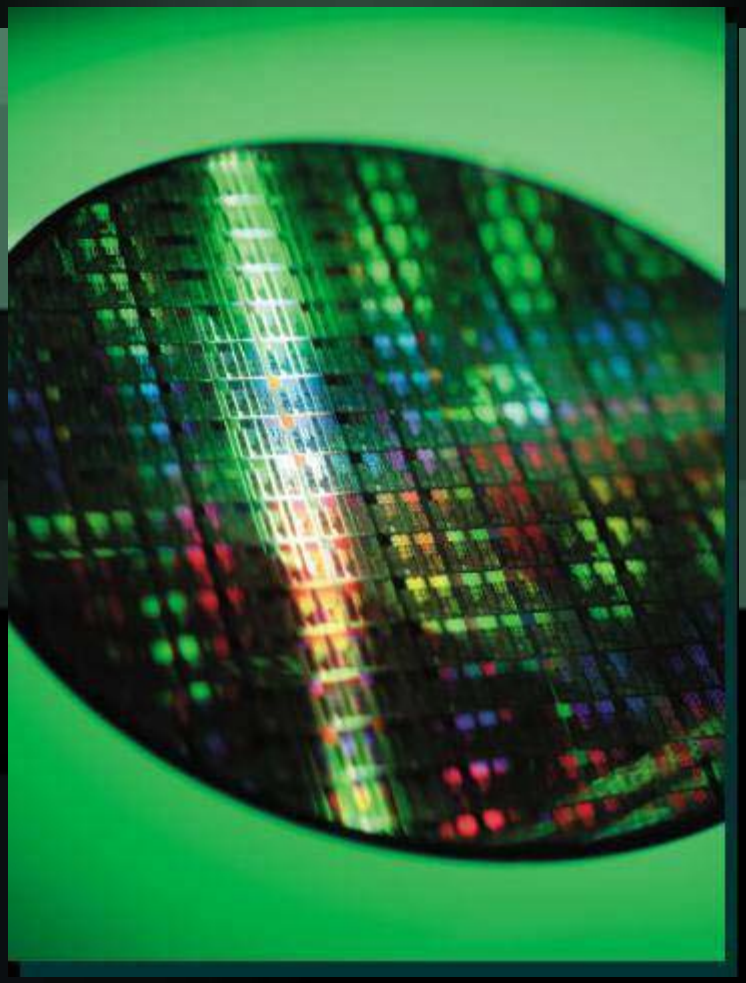
REX Prefix Byte

- Optional REX prefix specifies 64-bit operation size override and 3 additional register encoding bits
  - Extra registers encoded without altering existing instruction format
  - REX is actually a family of 16 prefixes (40-4F)
  - 64-bit mode Average instruction length increased by 0.4 bytes



Industry Transition to AMD64

- AMD Opteron™ and AMD Athlon™ 64 processors include AMD64 technology
- Transition to 64-bit computing will occur at the pace of demand for its benefits
- Transition from 286 to 386 is the perfect analogy
  - 386 was an initiative to create 32-bit capable processors
  - Initial users enjoyed highest performance 16-bit execution
  - Operating system and application development took time
  - Operating system support allowed 16-bit and 32-bit processes to co-exist and interoperate
  - 32-bit software is now the norm
  - Although the 386 was introduced in 1985, 16-bit compatibility was important for years
- Great compatibility combined with great performance is the only practical approach to introducing new capabilities



AMD64 Building Blocks

AMD64 Building Blocks

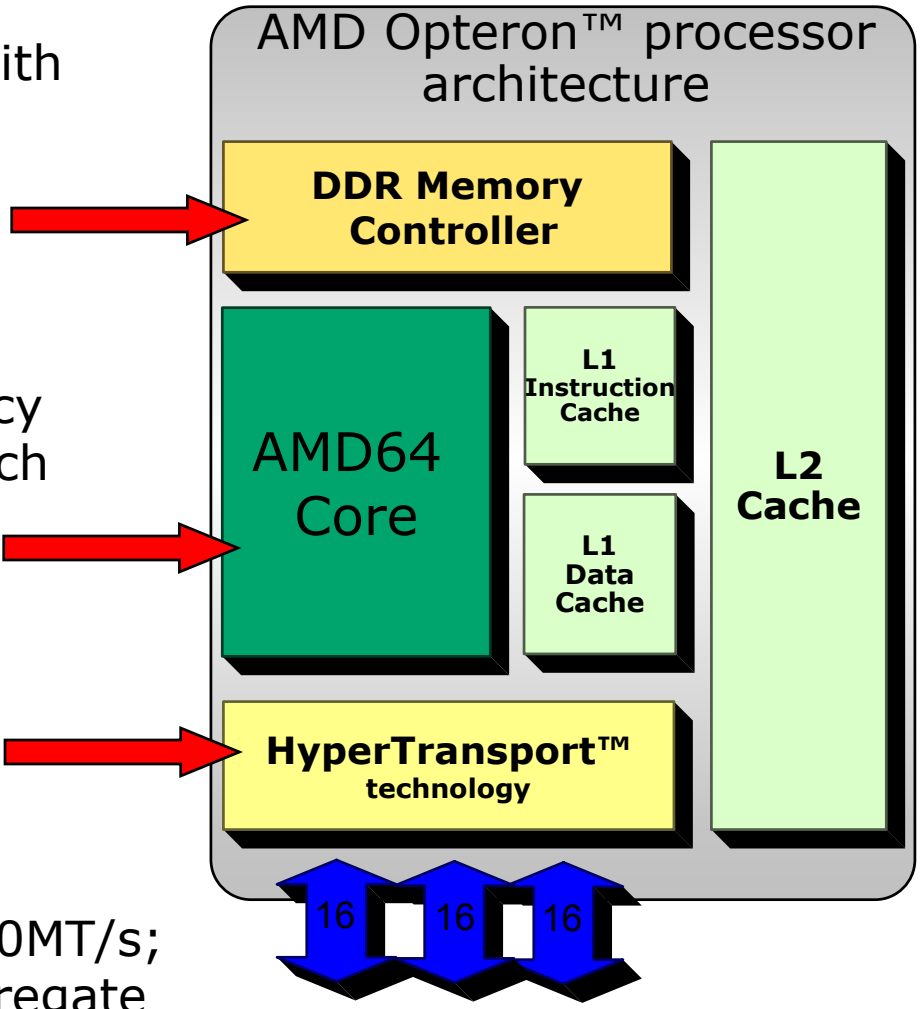
- Scalable systems must be built around efficient components
  - Power, cooling, board space and cost are crucial in building these.

- AMD provides the key building blocks for scalable AMD64 platforms:
  - Glueless multiprocessing through integrated memory controller and North bridge on the AMD Opteron™ processor die
  - HyperTransport™ technology interconnect and devices (PCI-X, AGP-8x, etc)
  - Reference platform designs to provide concrete examples to our OEM partners
  - AMD64 thermal/mechanical solutions are designed to meet the demanding requirements of PC and 1U form factors

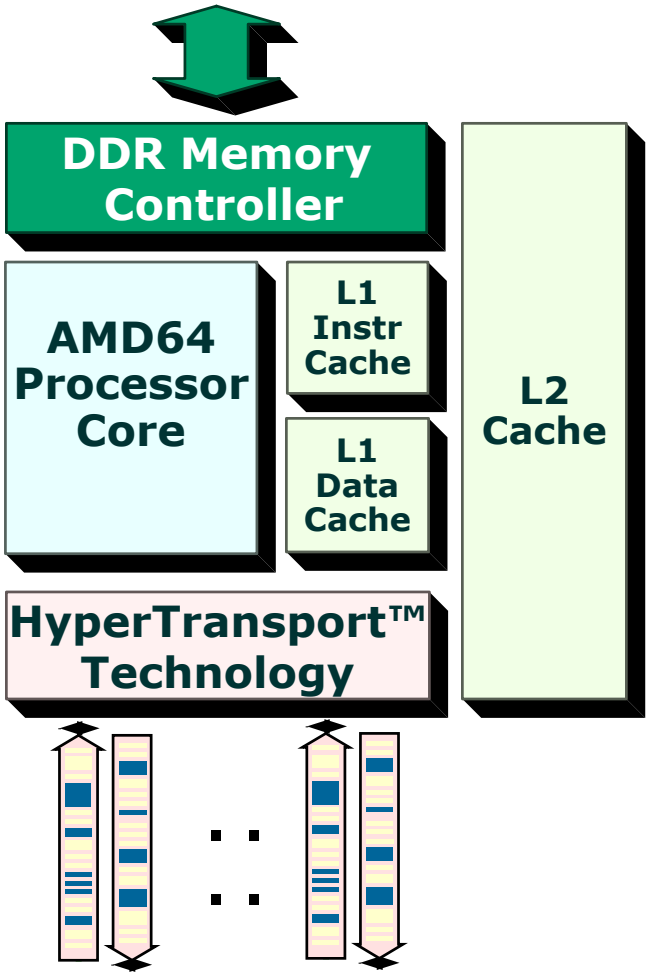
- An OEM, working with AMD, designs retail platforms customized to the OEM's needs and markets.

AMD Opteron™ Processor Overview

- ***Performance***
  - High-bandwidth integrated memory controller scales with processor frequency and number of processors
  - L2 1MB Cache
- ***Compatibility***
  - Approximately 10,000 legacy applications at time of launch
- ***Scalability***
  - Reduced costs for high-end systems
  - Removes I/O bottlenecks
  - Easy multiprocessor scaling
  - 16-bit HyperTransport™ technology links are at 1600MT/s; provides 6.4GB/s peak aggregate bandwidth

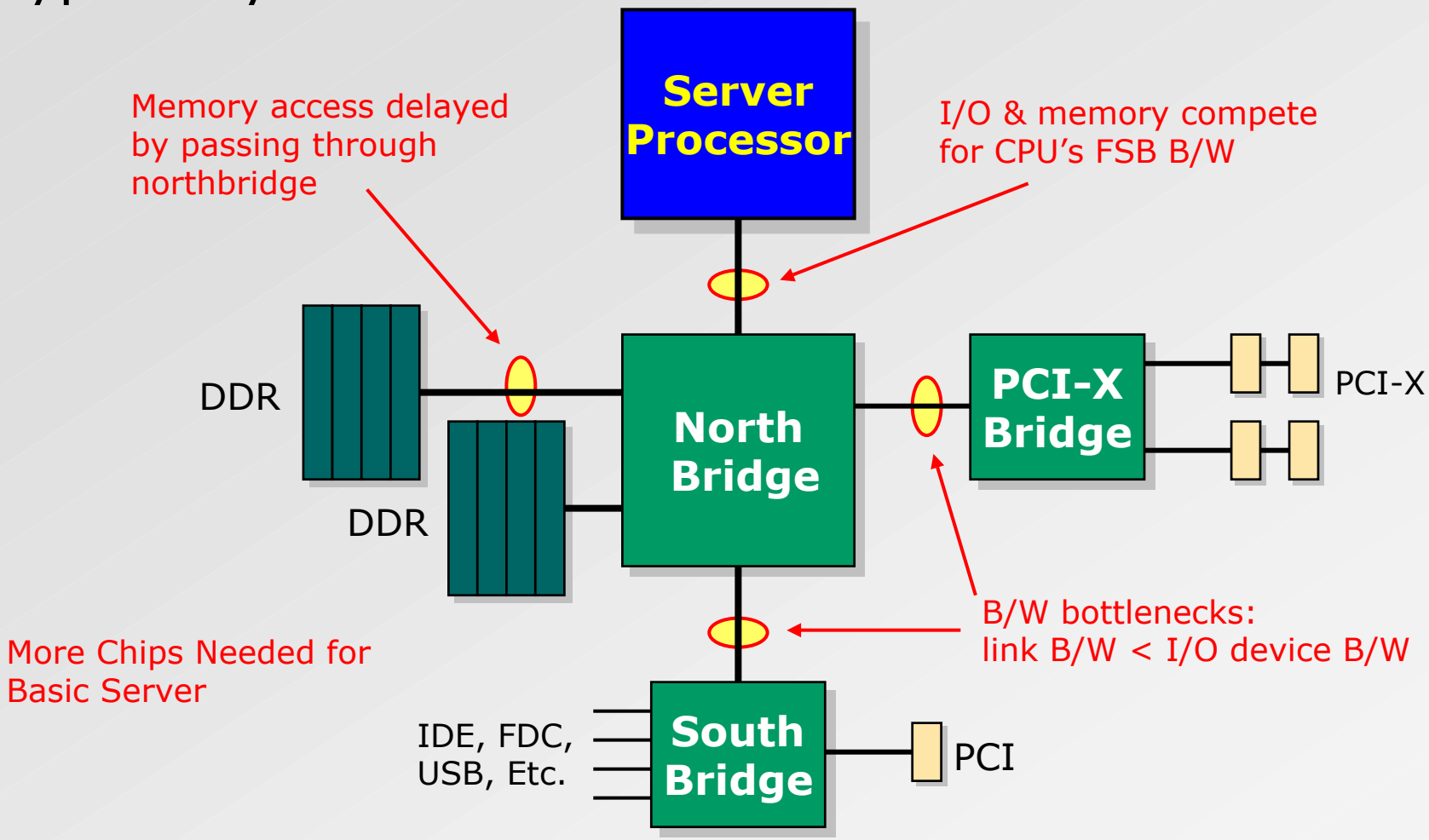


Integrated Memory Controller

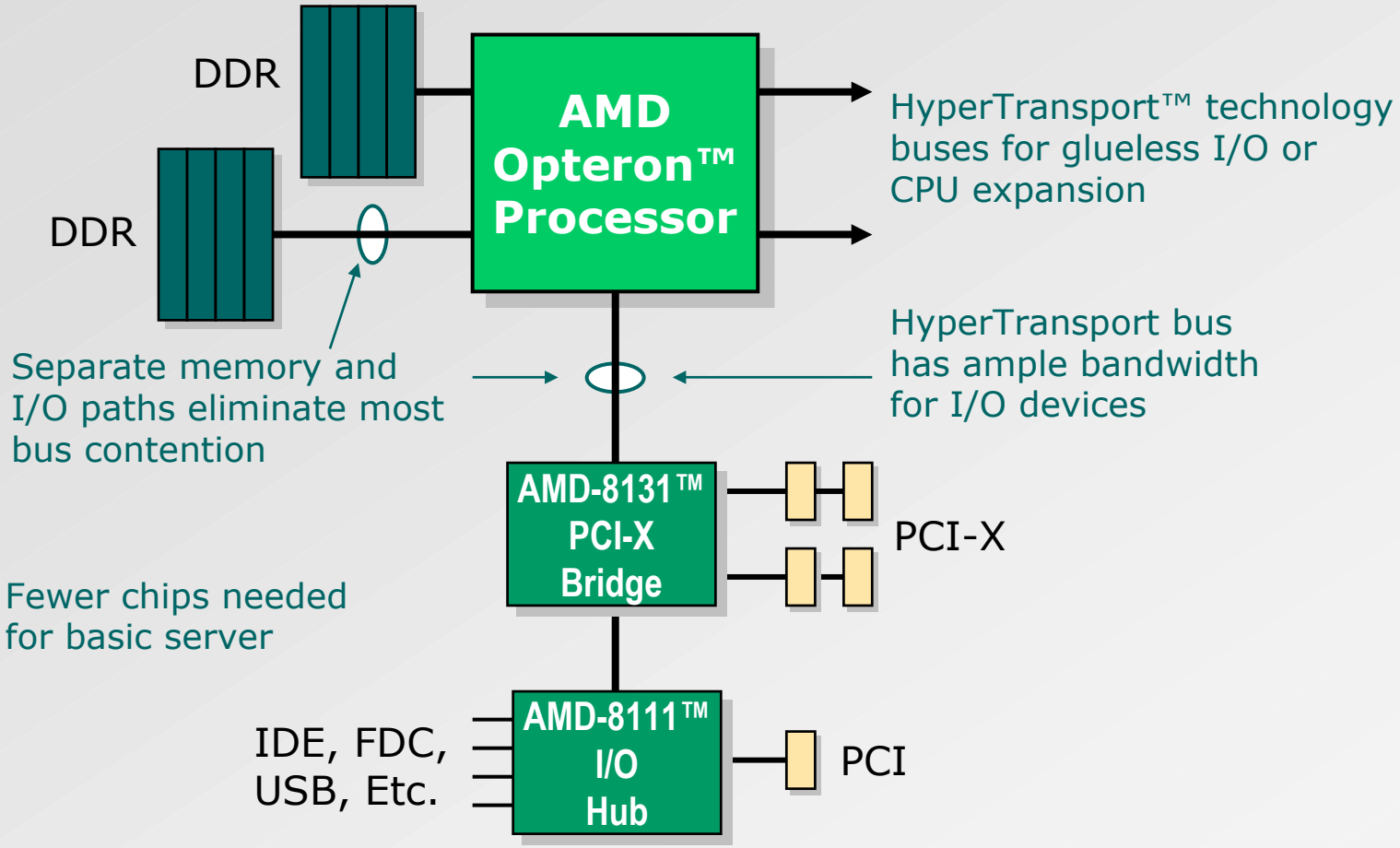


- Run memory controller at processor speeds rather than FSB speeds
  - Today's AMD Athlon™ XP processor north bridge memory controllers run at 133 MHz
- Dramatically decrease latency
  - QuantiSpeed™ architecture achieves ~160 ns best case latency
  - AMD64 architecture designed to achieve ~80 ns best case latency
  - Latency generally decreases further as the core frequency increases
- Add intelligence without decreasing performance
- Supports variety of DDR memories
  - 200, 266 and 333 MHz
  - Registered and unbuffered DIMMs
  - Future processor cores planned to support DDR-II, etc.

Typical System

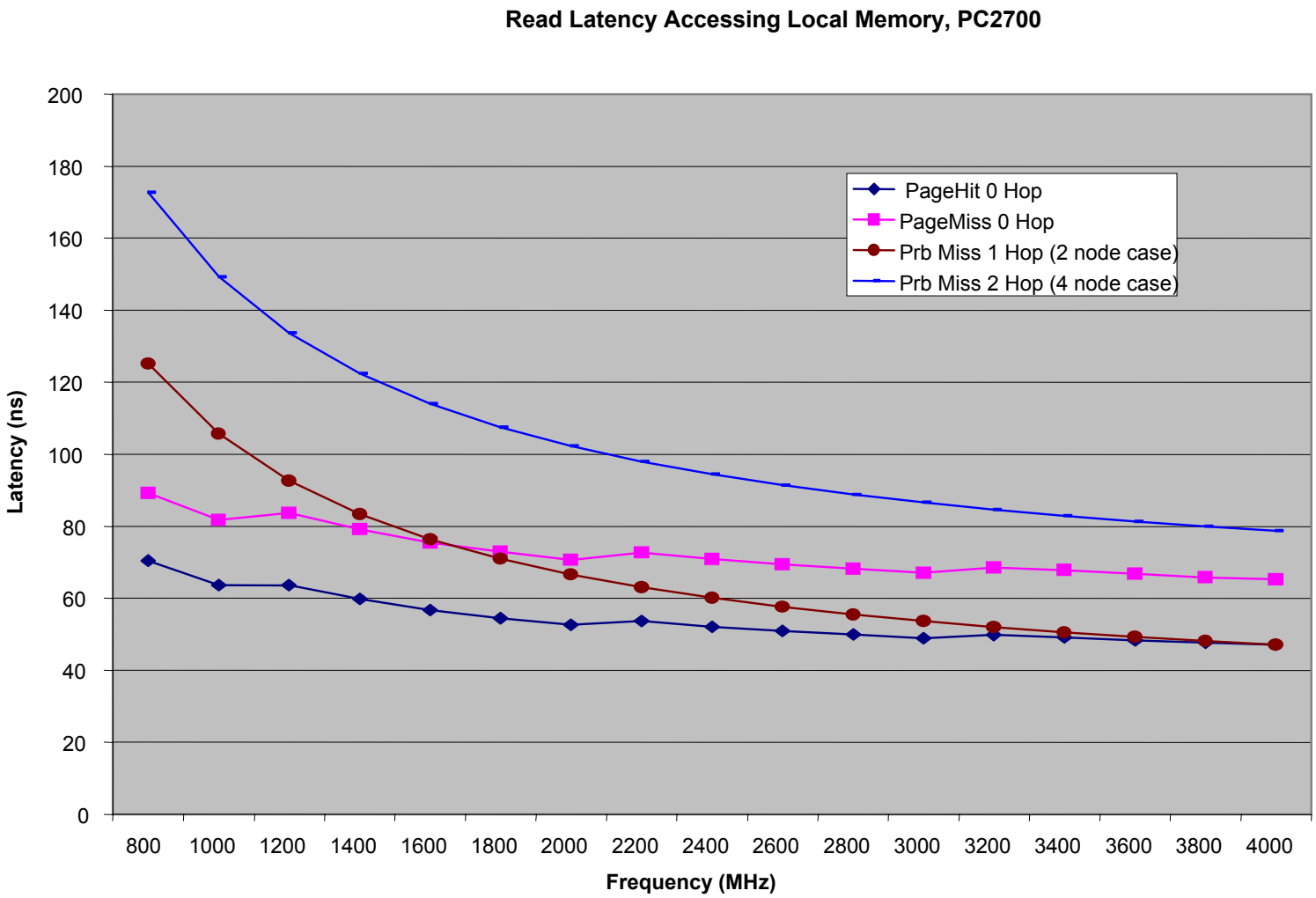


System Architecture

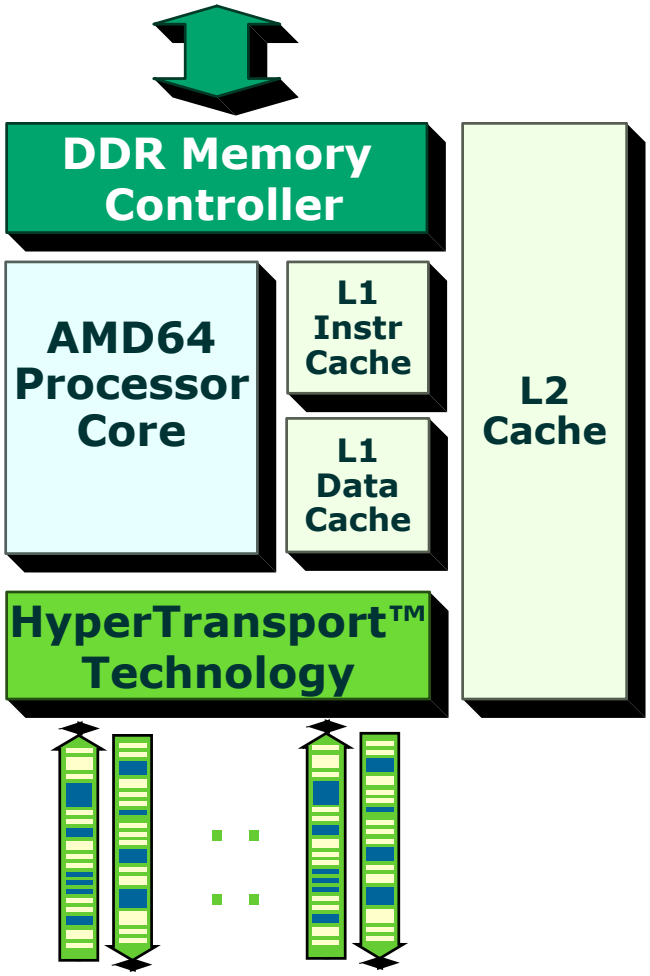


Integrated Memory Controller Latency

(Local memory access, registered memory, CAS2.5)



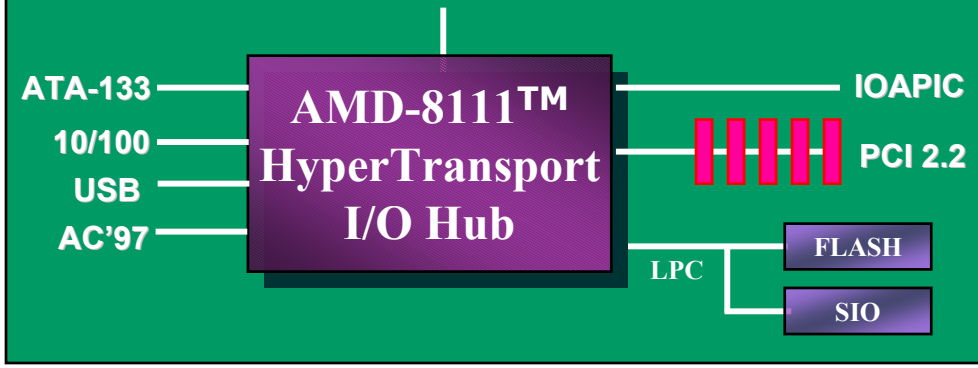
AMD64 Processors: HyperTransport™



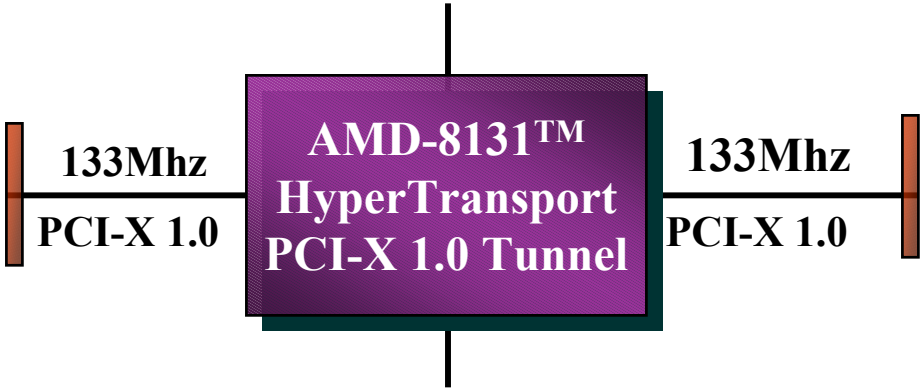
W = 2, 4, 8, 16, or 32-bits each way

- AMD developed HyperTransport™ technology
  - High-speed, low pin-count, asynchronous, chip-to-chip board level interconnect
  - Proven, industry-standard technology in production today – (Xbox)
- HyperTransport technology is not …
  - A replacement for PCI, PCI-X, PCI-Express
  - A networking fabric
- HyperTransport technology physical interface
  - Point-to-point, differential, low-voltage swing
  - HyperTransport 1.0 -> Up to 1600MT/s
  - HyperTransport 2.0 -> Beyond 4000MT/s
- HyperTransport technology logical interface
  - 100% PCI-compliant API
  - OS I/O (PCI) enumeration code untouched for AMD64 processor-based systems
- AMD made HyperTransport an OPEN STANDARD – High-profile, best-in-class partners
  - Broadcom, Cisco, NVIDIA, Sun, many more
  - www.hypertransport.org

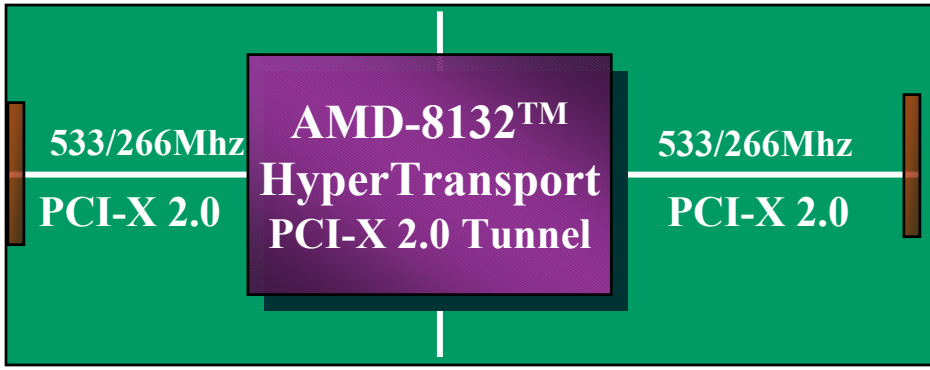
HyperTransport™ Technology



- HyperTransport™ technology I/O Hub
  - 8-bit host interface (800MB/s)
- Integrated I/O features
  - 10/100 ethernet controller
  - EIDE controller (supporting up to ATA-133)
  - PCI 2.2, LPC, USB, SMbus, IOAPIC, etc

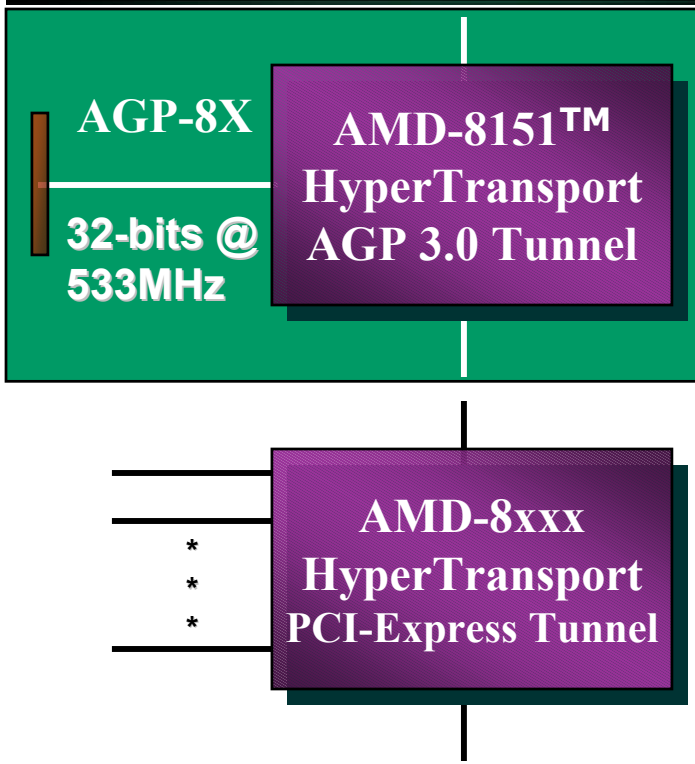


- HyperTransport technology tunnel
  - 16-bit host interface (6.4GB/s)
  - 8-bit next device interface (3.2GB/s)
- 2 independent PCI-X 1.0 ports
  - Supporting 133MHz, 100MHz, 66MHz, and legacy-PCI modes
  - I/O APIC



- HyperTransport technology tunnel
  - 16-bit host interface (8.0GB/s)
  - 16-bit next device interface (8.0GB/s)
- 2 independent PCI-X 2.0 Ports
  - Supporting 533MHz, 266MHz, 133MHz, 100MHz, 66MHz, and legacy-PCI modes
  - I/O APIC

HyperTransport™ Technology

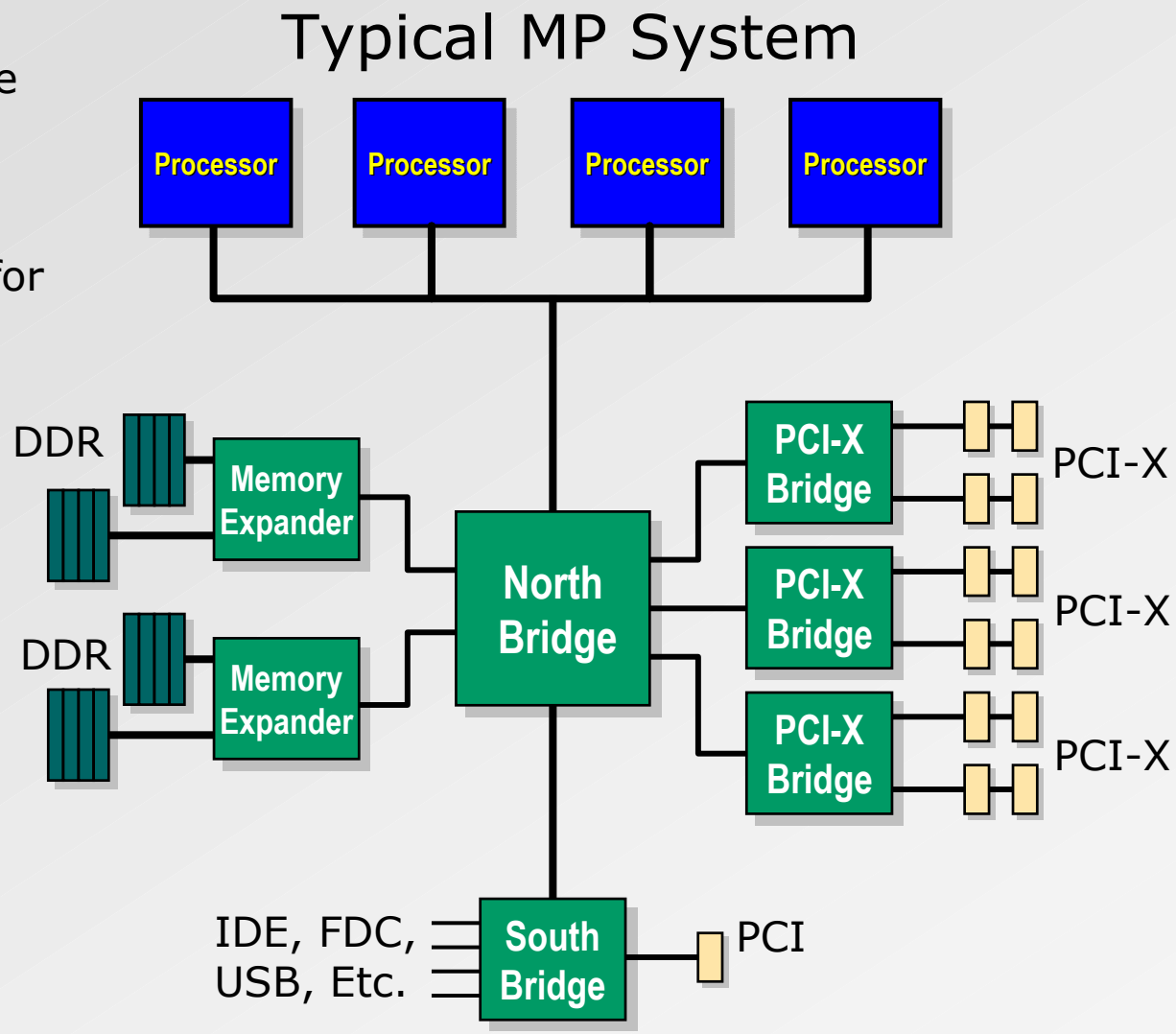


- HyperTransport™ technology tunnel
  - 16-bit host interface (6.4GB/s)
  - 8-bit next device Interface (3.2GB/s)
- AGP 3.0 Interface
  - AGP 3.0 (AGP-8X) & AGP 2.0 compliant interface.

- HyperTransport technology tunnel
  - 16-bit host interface
  - 16-bit next device interface
- PCI-Express ports
  - Several configurable primary Ports
  - Ports are sub dividable into several smaller ports

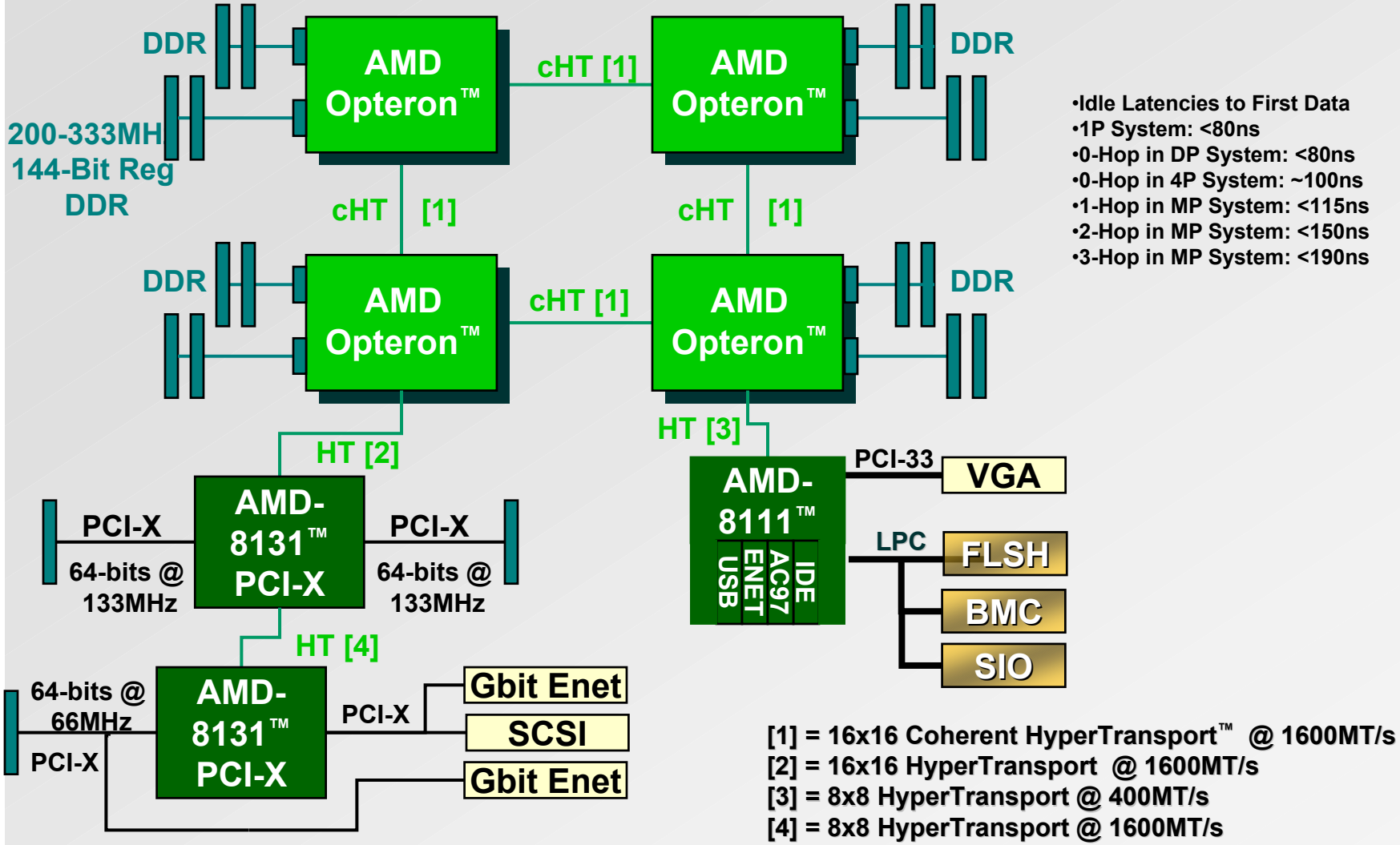
Typical Multiprocessing System

- System scalability limited by northbridge
  - Max of 4 processors
  - Processors compete for FSB bandwidth
  - Memory size and bandwidth are limited
  - Max of 3 PCI-X bridges
  - Many more chips required



800-Series 4P AMD Opteron™

Processor-based Server

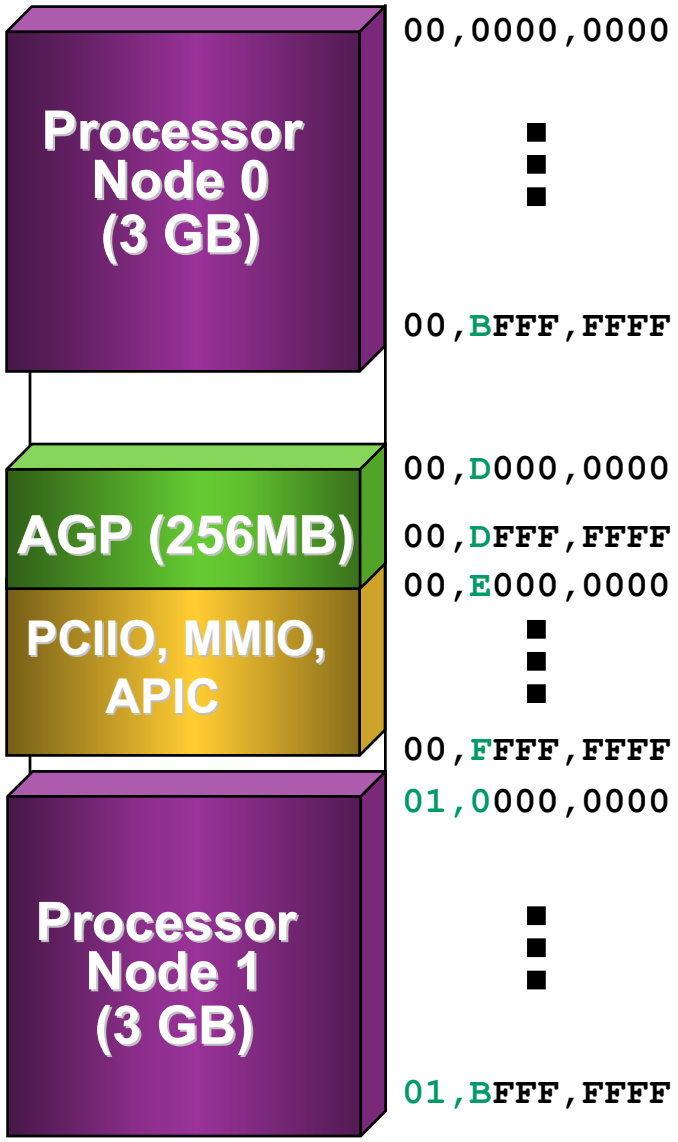


- Idle Latencies to First Data
- 1P System: <80ns
- 0-Hop in DP System: <80ns
- 0-Hop in 4P System: ~100ns
- 1-Hop in MP System: <115ns
- 2-Hop in MP System: <150ns
- 3-Hop in MP System: <190ns

[1] = 16x16 Coherent HyperTransport™ @ 1600MT/s
[2] = 16x16 HyperTransport @ 1600MT/s
[3] = 8x8 HyperTransport @ 400MT/s
[4] = 8x8 HyperTransport @ 1600MT/s

Physical Memory Map Layout

- Assume
  - 2-CPU WS, 3GB DRAM per CPU
  - Non-node interleaved DRAM map
- BIOS maps these below 4GB so legacy OS & PCI devices can access:
  - AGP Aperture, PCI I/O-Map region, PCI Mem-Map region, APIC
- Simple mapping places 2nd CPU's memory above 4GB:
  - Other optimizations possible.



Simple SW Model for NUMA

- NUMA bring dramatic scalability advantages
  - But software management is hard to get right
- SMP systems bring dramatically simplified software model
  - But memory system doesn't scale up as you add processors
- AMD Opteron™ processor provides benefits of both:
  SMP view for Software
  - Physical address space is flat and fully coherent
  - Latency difference between local and remote memory in an 8P system is comparable to the difference between a DRAM page hit and page miss
  - DRAM can be contiguous or interleaved
- MP support designed into processor & system from the beginning
  - Lower overall system chip count increases reliability and lowers cost
  - All MP system functions use CPU technology and frequency
- Latency shrinks quickly by increasing CPU and HyperTransport™ technology link speed
  - Additional processor nodes bring increased memory bandwidth and great overall system throughput


AMD
AMD64 Systems

AMD Opteron™ Processor Platforms

Platform	**Khepri (Newisys)**
Processor	2P AMD Opteron
Memory	DDR-333 x4 each
Physicals	1U
I/O	PCIX, 2xGig Ether, SCSI
Management	Proprietary Service Processor

Company	**TYAN**
Processor	2P AMD Opteron
Memory	DDR-333 x4 each (typical)
Physicals	2U or pedestal
I/O	PCIX, 2xGig Ether, SCSI
Management	IPMI 1.5/Remote Mgmt LAN

Company	**GIGABYTE**
Processor	2P AMD Opteron
Memory	DDR-333 x4 each (typical)
Physicals	2U or pedestal
I/O	PCIX, 2xGig Ether, SCSI
Management	IPMI 1.5/Remote Mgmt LAN

Platform	**Quartet**
Processor	4P AMD Opteron
Memory	DDR-333 x4 each
Physicals	4U N+1 PSU
I/O	PCIX, 2xGig Ether, SCSI
Management	IPMI 1.5/Remote Mgmt LAN

Company	**ARIMA**
Processor	2P AMD Opteron
Memory	DDR-333 x4 each (typical)
Physicals	2U or pedestal
I/O	PCIX, 2xGig Ether, SCSI
Management	IPMI 1.5/Remote Mgmt LAN

Motherboard	**Rhapsody (RDK)**
Processor	2P AMD Opteron
Memory	DDR-333 x4 each (typical)
Physicals	2U or pedestal
I/O	PCIX, 2xGig Ether, SCSI
Management	IPMI 1.5/Remote Mgmt LAN

Company	**MSI**
Processor	2P AMD Opteron
Memory	DDR-333 x4 each (typical)
Physicals	2U or pedestal
I/O	PCIX, 2xGig Ether, SCSI
Management	IPMI 1.5/Remote Mgmt LAN

Available from OEMs world-wide

1H 2003

Manageability

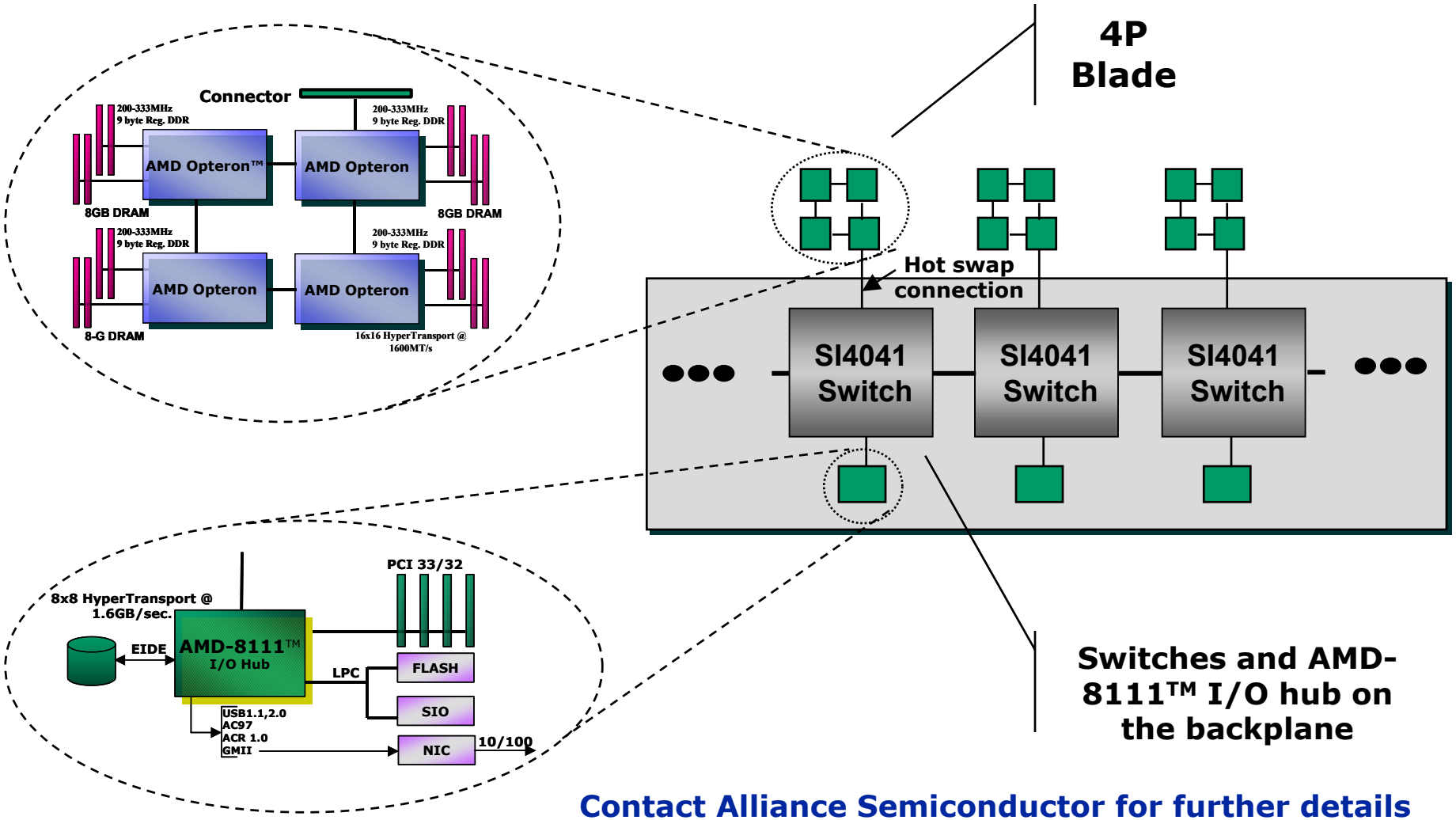
- AMD committed to industry standard manageability
- AMD reference designs and Validated Server Program (VSP) products provide standards based management
  - WBEM/CIM – Web-Based Enterprise Management
    - Initiative of the Distributed Management Task Force (DMTF) to provide standard framework for the for the management of clients, servers, networks, storage, etc…
  - IPMI – Intelligent Platform Management Interface
    - Standard for server management, hardware health monitoring and remote system control (power/reset/etc)
  - SMBIOS – System Management BIOS
  - PXE – Pre-boot eXecution Environment

IPMI

- OSA is the management solution provider for AMD processor-based server reference platforms and VSP products
  - Supports Quartet and Serenade platforms
  - Base IPMI 1.5 support
  - Multi-platform management and monitoring
  - Centralized event and alert handling
  - In-band and out-of-band management
  - 3-tiered architecture to support internet and multi-console access
  - SSL security

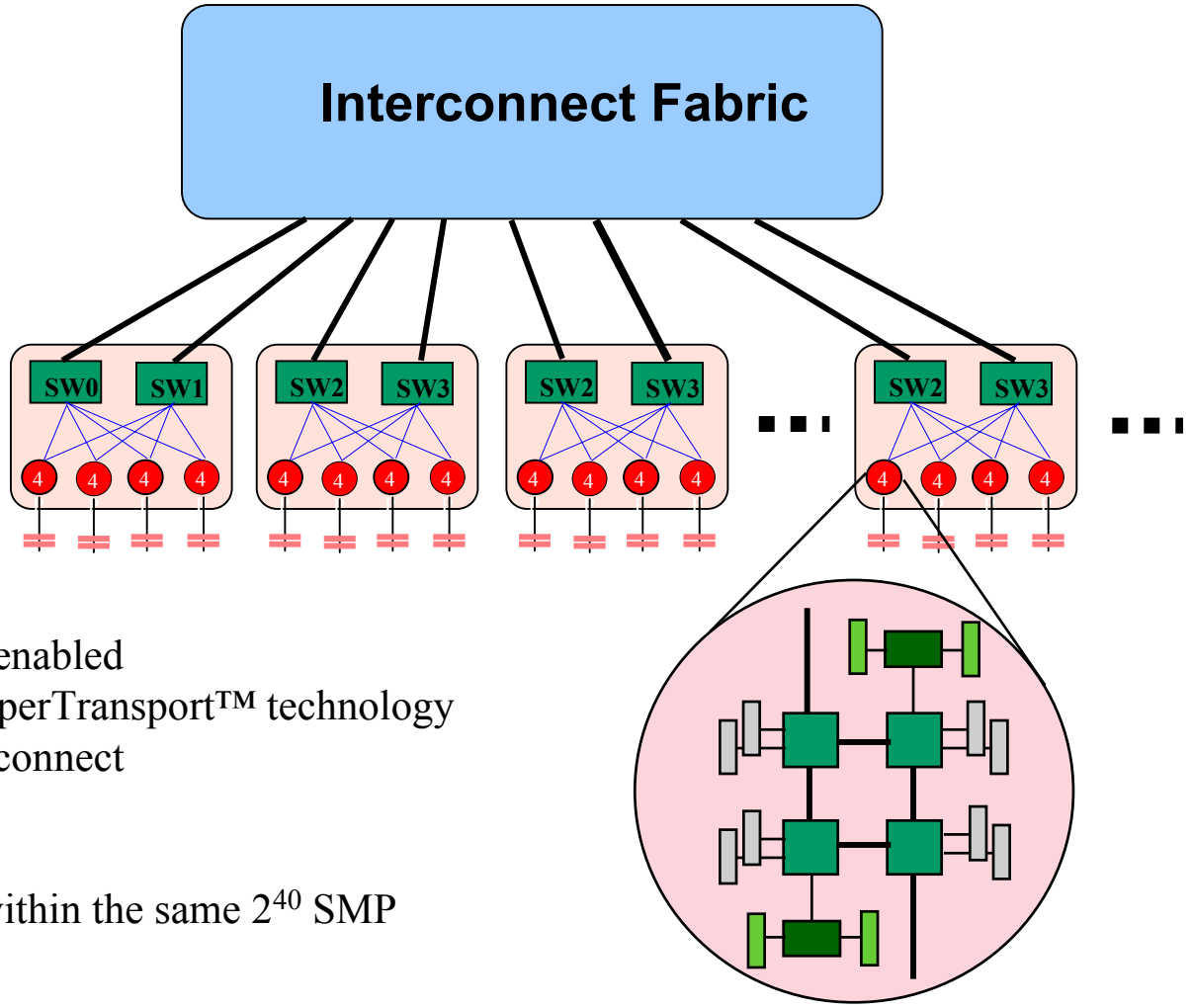
HyperTransport™ Technology Backplane

- **Using non-coherent interconnect**



Contact Alliance Semiconductor for further details

Scaling Beyond 8 Processors



- Scaling beyond 8P is enabled
- External Coherent HyperTransport™ technology switch Coherent Interconnect
  - ✓ Snoop filter
  - ✓ Data caching
- Up to 16 processors within the same 2^{40} SMP memory space

AMD

AMD64 Software

32-bit OS & Application Support

- 32-bit software does not have to be ported since the AMD Opteron™ processor is compatible with 32-bit OS, applications, and drivers
  - Natively supports x86 instruction set
  - Floating-point model is x87
  - SSE, SSE2, MMX, 3DNow!, and in-line ASM are supported
- AMD Opteron processor offers a high-performance platform for 32-bit software
  - Takes full advantage of core enhancements offered by AMD Opteron processor
  - Is expected to progressively run faster as systems speed up
  - AMD Opteron processor-based systems are posting leading 32-bit benchmarks, including SPECint_rate, SPECfp®_rate, SPECWEB99, MMB2, and TPC-C

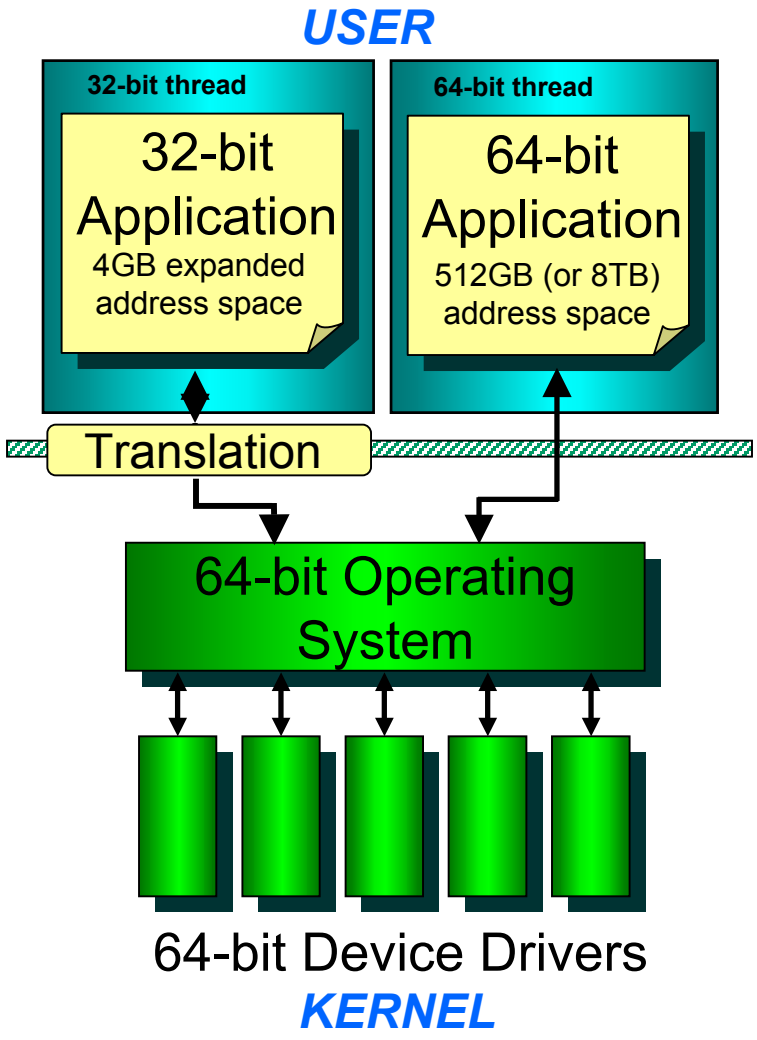
64-bit OS & Application Interaction

32-bit Compatibility Mode

- 64-bit OS runs existing 32-bit APPs with leading edge performance
- No recompile required, 32-bit code directly executed by CPU
- 64-bit OS provides 32-bit libraries and "thunking" translation layer for 32-bit system calls.

64-bit Mode

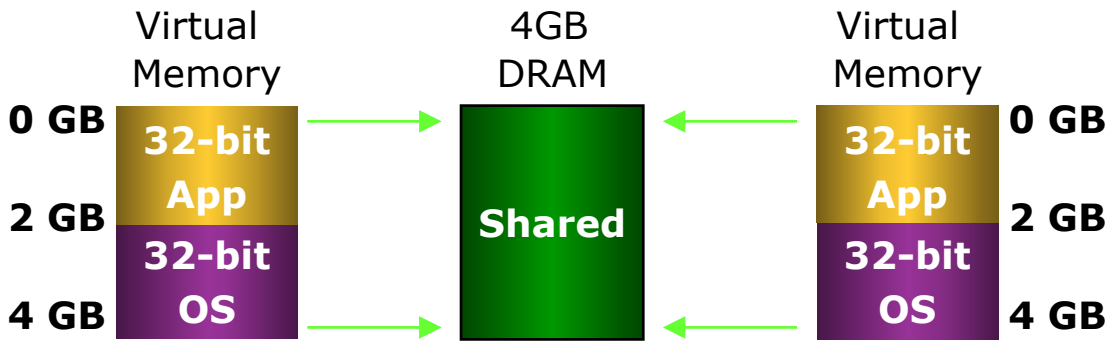
- Migrate only where warranted, and at the user's pace to fully exploit AMD64
- 64-bit OS requires all kernel-level programs & drivers to be ported.
- Any program that is linked or plugged in to a 64-bit program (ABI-level) must be ported to 64-bits.



Increased Memory for 32-bit Applications

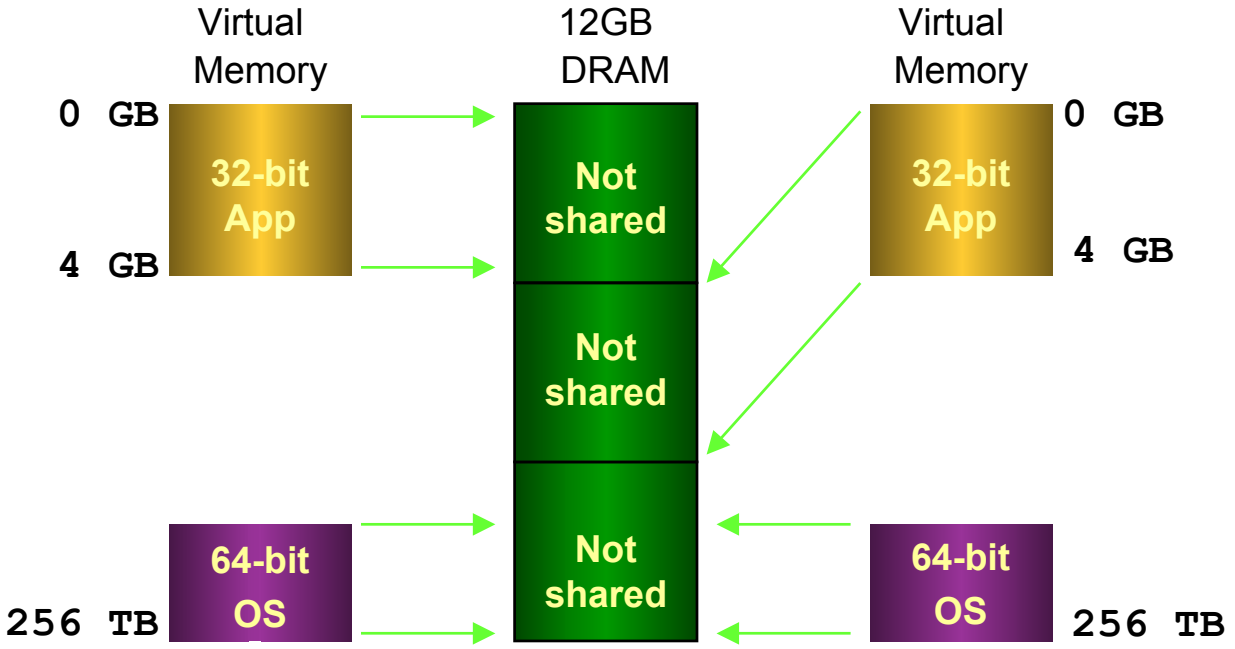
32-bit server, 4 GB DRAM

- OS & App share small 32-bit VM space
- 32-bit OS & applications all share 4GB DRAM
- Leads to small dataset sizes & lots of paging



64-bit server, 12 GB DRAM

- App has exclusive use of 32-bit VM space
- 64-bit OS can allocate each application large dedicated portions of 12GB DRAM
- OS uses VM space way above 32-bits
- Leads to larger dataset sizes & reduced paging



Current Operating System Support

	Operating System	Type
	SuSE Linux Enterprise Server (SLES) 8	32 & 64-bit
	SuSE Linux 8.2 Personal & Professional	32-bit
	UnitedLinux Version 1.0 code base by UnitedLinux Consortium	32 & 64-bit
	Conectiva Linux Enterprise Edition	32-bit
	Linux AMD64 kernel patches	64-bit
	Mandrake Linux 9.0	64-bit
	Mandrake Linux 9.1	32-bit
	Mandrake Linux Corporate Server 2.1	32-bit
	NetBSD	32 & 64-bit
	SCO Linux	32-bit
	Scyld Beowulf Cluster Operating System	32-bit
	Solaris 9 for x86 (32-bit)	32-bit
	Turbolinux 8 for AMD64	32 & 64-bit
	Windows® 2000 Server (32 bit)	32-bit
	Windows Server 2003 (32-bit)	32-bit

Planned Operating Systems Support

Operating System	Type	Available
Red Hat Enterprise Linux AS 3.0	32- & 64-bit	Beta mid-2003
SuSE Linux 8.3 Personal & Professional	32- & 64-bit	September 2003
Windows® XP Professional for AMD64	64-bit	Beta mid-2003
Windows Server 2003 for AMD64	64-bit	Beta mid-2003

Support for AMD's 64-bit processors is scheduled for Windows Server 2003 Service Pack 1, "hopefully by the end of the year." David Thompson, Microsoft VP

http://www.extremetech.com/article2/0,3973,1061703,00.asp

Linux on the AMD Opteron™ Processor is a port to enhanced x86 architecture -- one that delivers 64-bit punch, while maintaining compatibility for classic x86 32-bit applications.

http://newsforge.com/newsforge/03/04/21/1914216.shtml?tid=7

AMD64 Developer Tools

- GNU compilers
  - GCC 3.2.2 - 32-bit and 64-bit
  - GCC 3.3 - optimized 64-bit

Optimized Compilers Are Reaching Production Quality

1.8 MHz AMD Opteron Processor– SPECint2000

Compiler	OS	Base	Peak
Intel C/C++ 7.0	Windows Server 2003	1095	1170
Intel C/C++ 7.0	Linux/x86-64	1081	1108
Intel C/C++ 7.0	Linux (32-bit)	1062	1100
GCC 3.3 (64-bit)	Linux/x86-64	1045	
GCC 3.3 (32-bit)	Linux/x86-64	980	
GCC 3.3 (32-bit)	Linux (32-bit)	960	

http://www.aceshardware.com/

- **PGI Workstation 5.0 beta**
  - Optimized Fortran 77/90, C/C++ for 32-bit Linux and Windows and 64-bit Linux
  - Product release by end of June
  - Performance goals are to be on parity with Intel compilers

- **Visual C, C++ for AMD64 is included with Windows® for AMD64 alpha release and current build reflects initial optimization**

FORTRAN and C/C++ Compiler Support

AMD

- AMD and STMicroelectronics are working together to bring The Portland Group Compiler Technology to AMD64
  - Support will include
    - F90 & F77
      - Some F95 extensions also included
    - Optimized 32-bit and 64-bit code generation
    - Both Linux and Windows® will be supported
    - OpenMP support
    - Full debugging support
  - Compiler is designed to provide performance that can meet or exceed that of competitive compilers in the market today
- Beta versions freely available today at www.pgroup.com/AMD64
- Commercial release of Workstation 5.0 is scheduled for 6/30/03 – CDK in July

The Portland Group
Compiler Technology

STMicroelectronics
More Intelligent Solutions

PGI CDK 5.0 - Additional Features

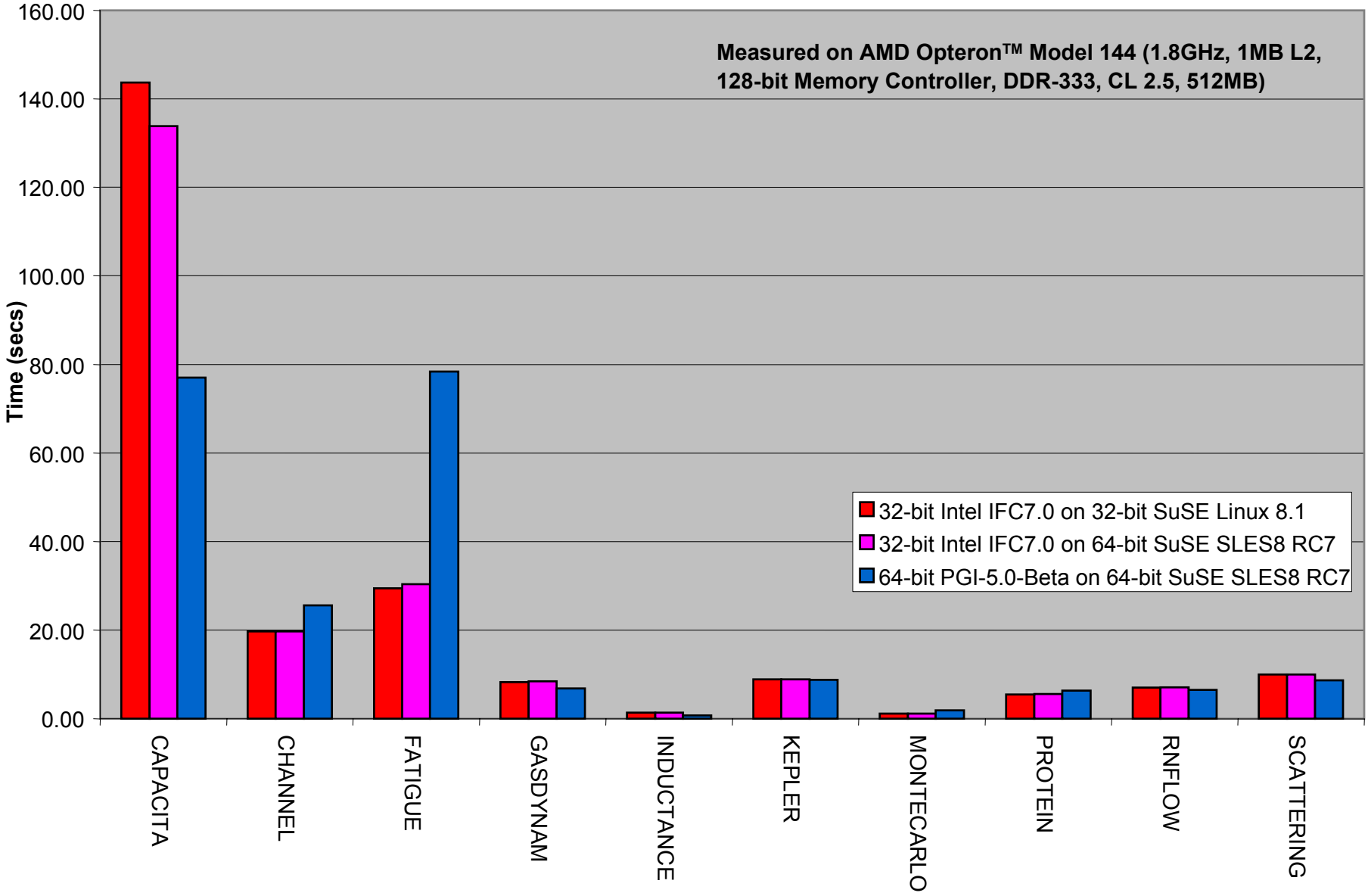
- Tentative Availability - July, 2003
- Highly Scaleable
- MPI-CH - Pre-configured libraries and utilities for ethernet-based x86 and AMD64/Linux clusters
- PBS – Portable Batch System batch-queuing from NASA Ames and MRJ Technologies
- ScaLAPACK - Pre-compiled distributed-memory parallel Math Library
- ACML – The AMD Core Math Library is planned to be included
- Training – Tutorials (OSC), exercises, examples and benchmarks for MPI, OpenMP and HPF programming

Optimized Numerical Libraries: ACML

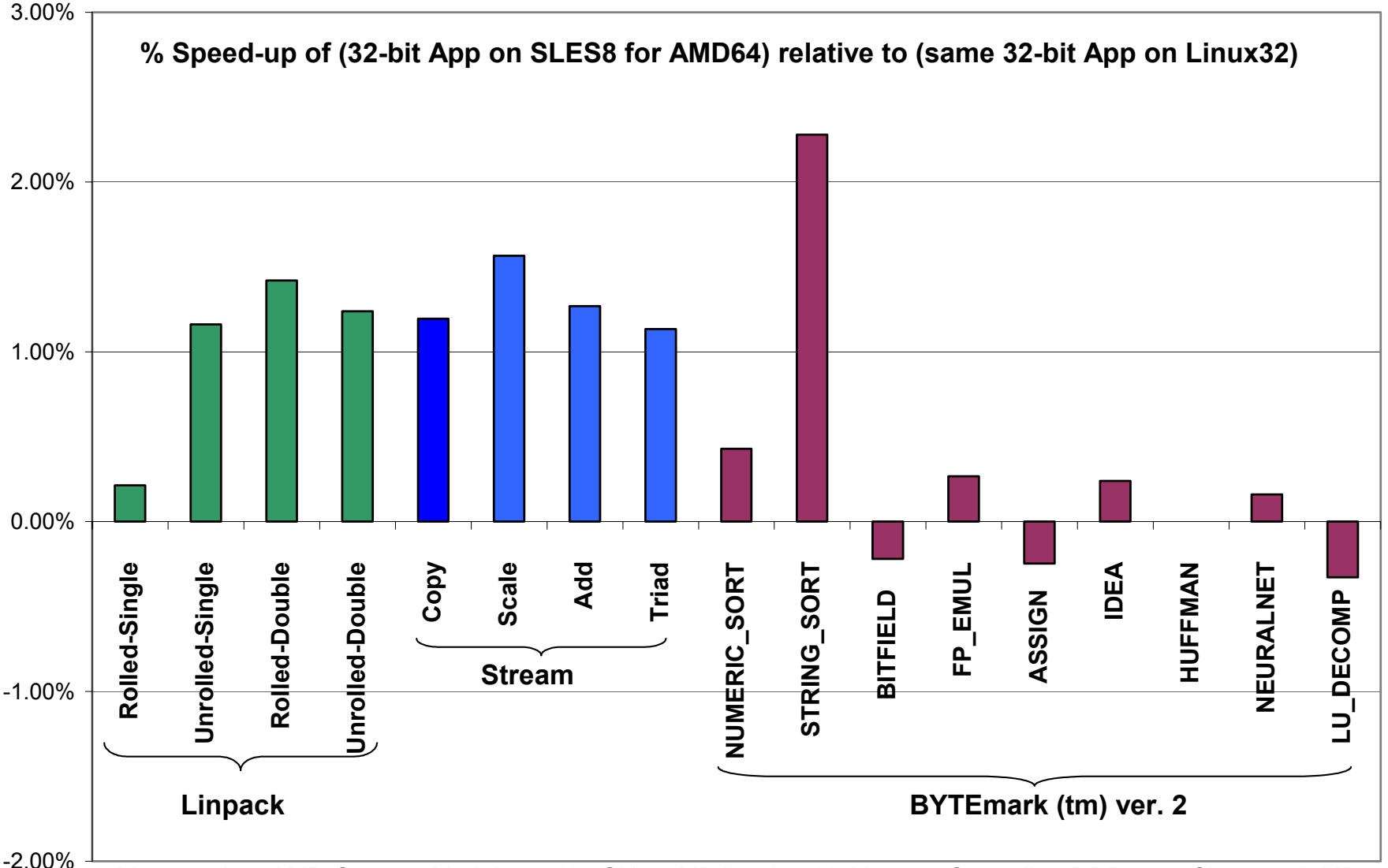
- AMD and The Numerical Algorithms Group (NAG) are jointly developing the AMD Core Math Library (ACML)
  - For use with mathematical, engineering, scientific and financial applications as well as general HPC computing
- ACML is comprised of:
  - Basic Linear Algebra Subroutines (BLAS) levels 1, 2 and 3
  - A wide variety of Fast Fourier Transforms (FFTs)
  - Linear Algebra Package (LAPACK)
- ACML will have the following features:
  - Fortran and C Interfaces
  - Highly optimized routines for the AMD64 Instruction Set
  - Ability to address single-, double-, single-complex and double-complex data types
  - Will be available for commercially available OSs
- ACML 1.0 is scheduled to be released on 6/30/03 and will be freely downloadable from www.developwithamd.com/acml

AMD64 Performance

Fortran Polyhedron Compiler Comparison

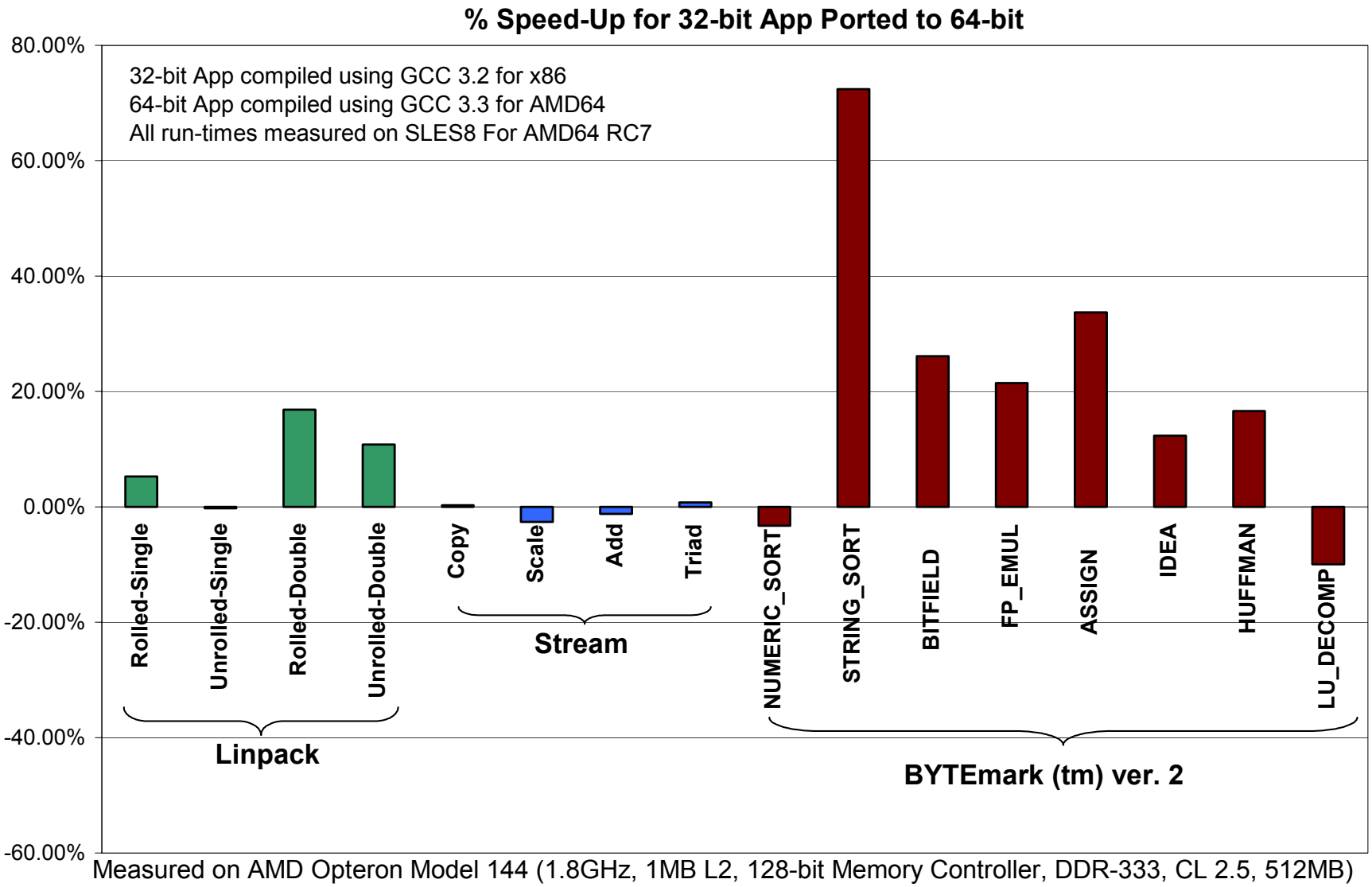


32-bit App Performance on 64-bit Linux



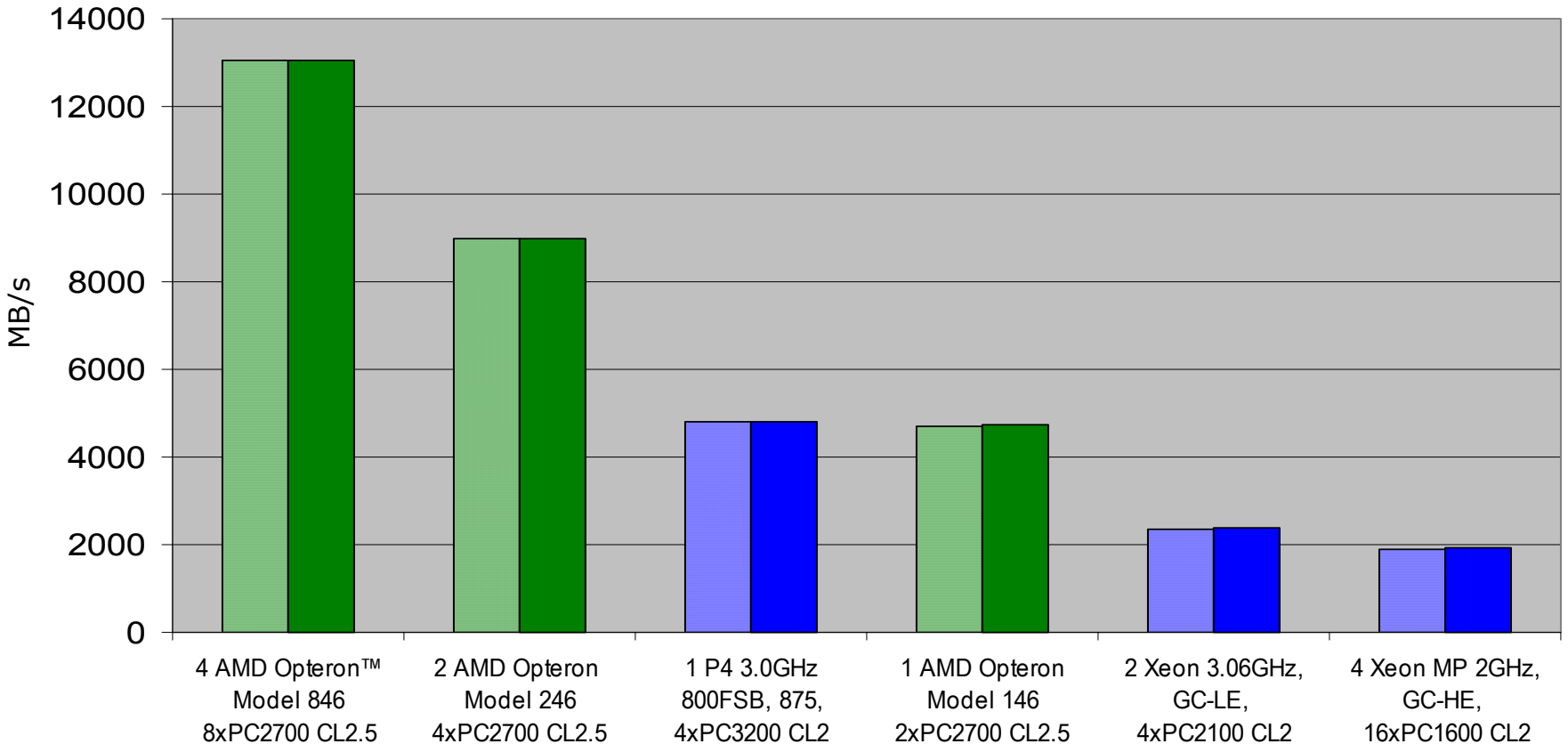
Measured on AMD Opteron Model 144 (1.8GHz, 1MB L2, 128-bit Memory Controller, DDR-333, CL 2.5, 512MB)

32-bit vs 64-bit App Performance



Measured on AMD Opteron Model 144 (1.8GHz, 1MB L2, 128-bit Memory Controller, DDR-333, CL 2.5, 512MB)

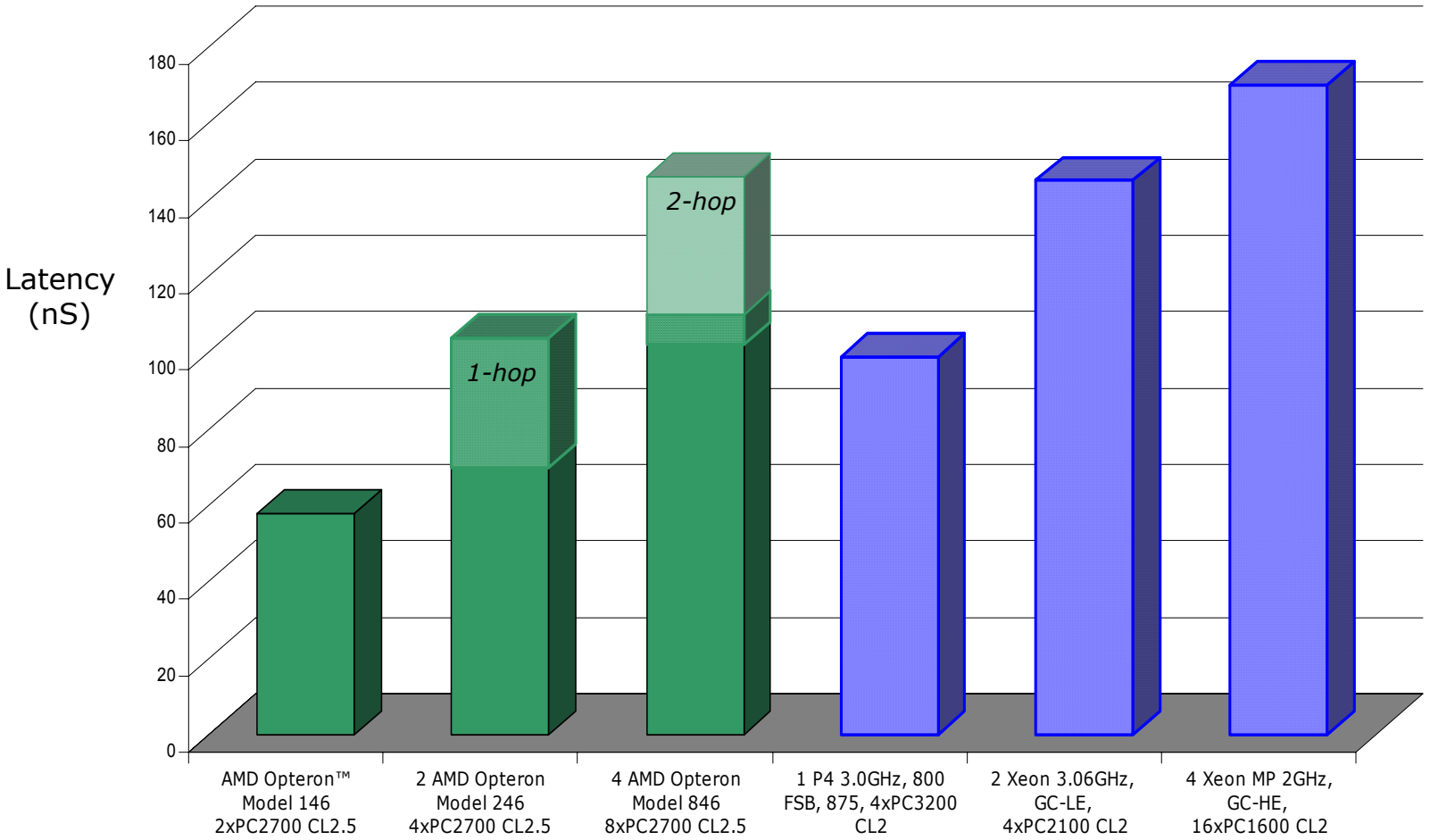
Scalable Memory Bandwidth
Sisoft Sandra Standard 2003



All benchmarks run on Microsoft® Windows® Server 2003 Enterprise Edition

Low Memory Latency

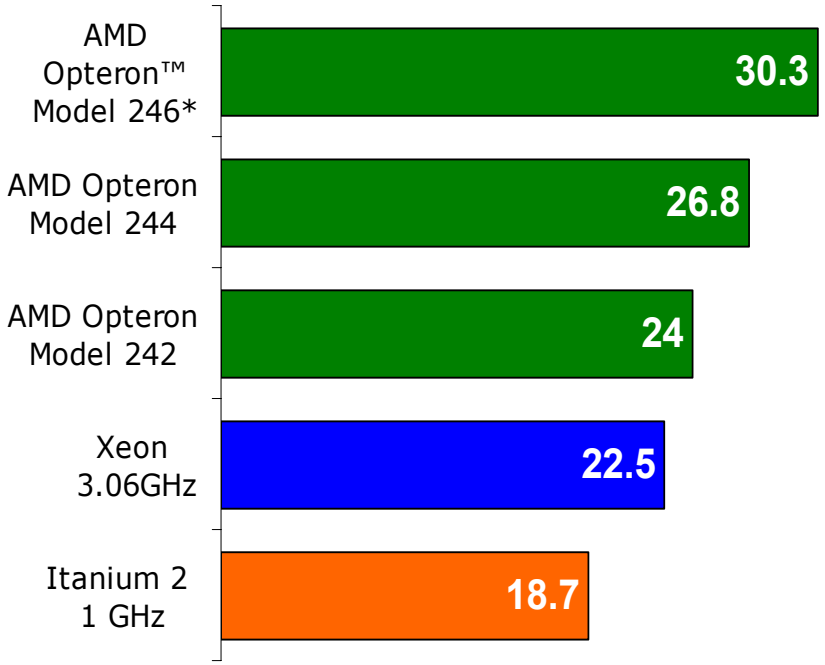
ScienceMark 2.0 Beta, 512-Byte Stride



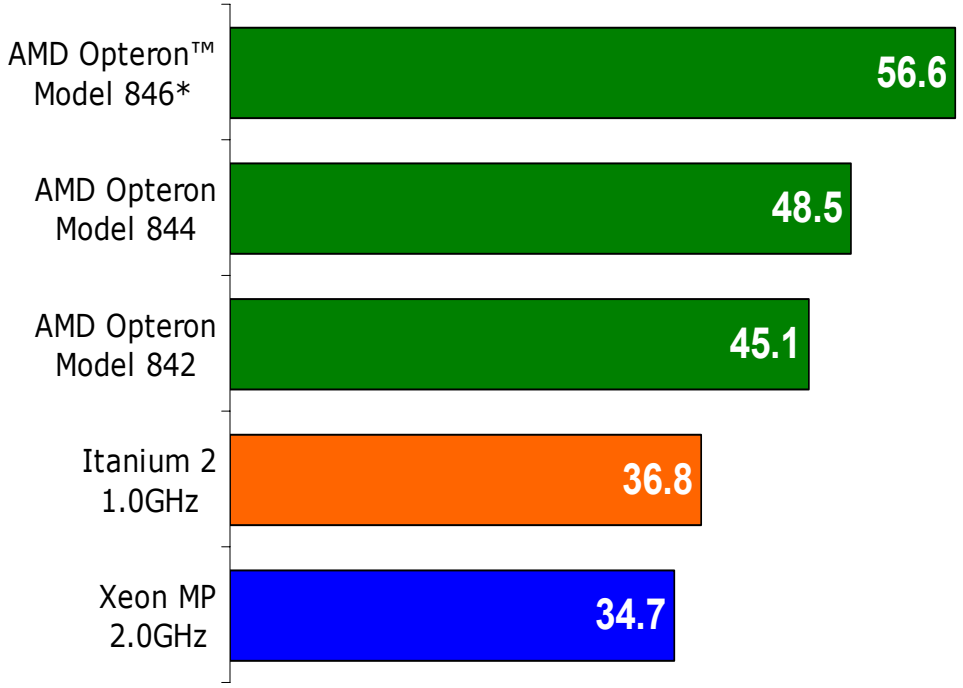
All benchmarks run on Microsoft® Windows® Server 2003 Enterprise Edition

MP Integer Performance

SPECint®_rate2000 Performance (Peak, 2P)



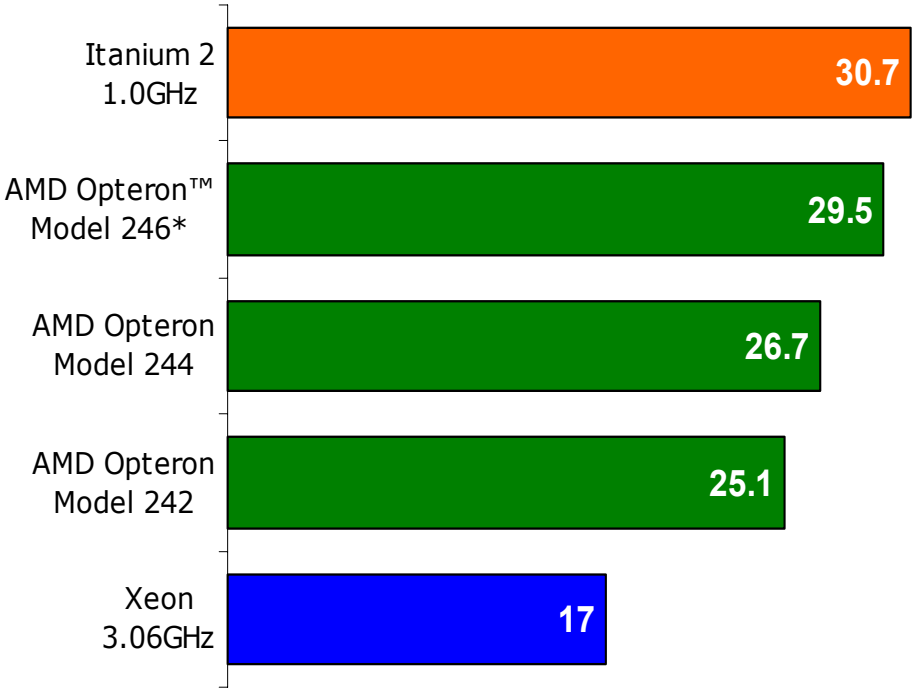
SPECint®_rate2000 Performance (Peak, 4P, Windows®)



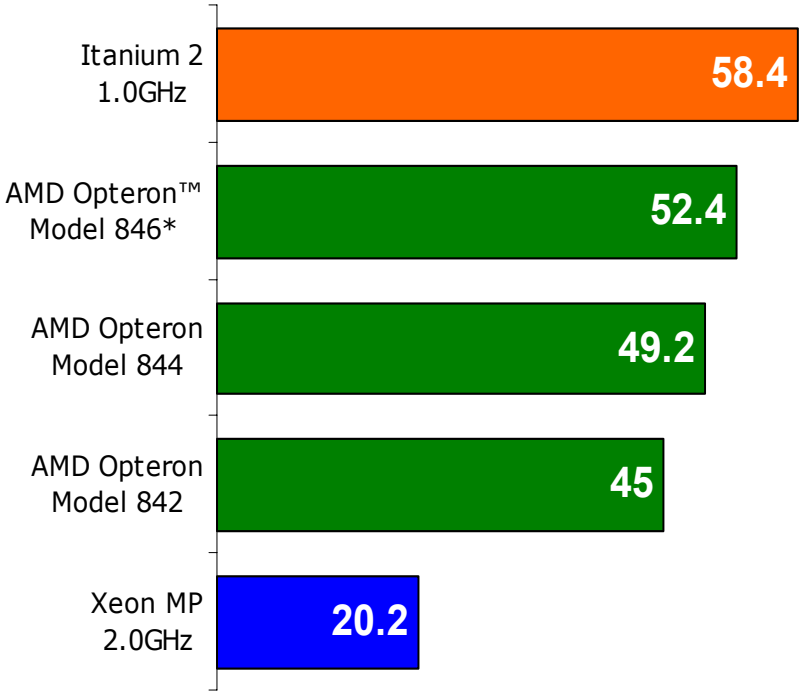
* 246 & 846 Results are estimated pending final SPEC submission
SPEC and the benchmark name SPECint are registered trademarks of the Standard Performance Evaluation Corp. Competitive numbers shown reflect results published on www.spec.org as of June 17, 2003. For the latest SPEC results visit <http://www.spec.org/>.

MP Floating-point Performance

SPECfp®_rate2000 Performance (Peak, 2P)



SPECfp®_rate2000 Performance (Peak, 4P)



* 246 & 846 Results are estimated pending final SPEC submission
SPEC and the benchmark name SPECfp are registered trademarks of the Standard Performance Evaluation Corp. Competitive numbers shown reflect results published on www.spec.org as of June 17, 2003. For the latest SPEC results visit <http://www.spec.org/>.

High Performance Linpack

GOTO Library Results AMD Opteron™ system	# P	Rmax (GFlops)	Nmax (order)	N1/2 (order)	Rpeak (GFlops)	GFLOP/ Proc	Rmax / Rpeak
4P AMD Opteron 1.8GHz 2GB/proc PC2700 8GB Total	**4**	**12.06**	**28000**	**1008**	**14.4**	**3.02**	**83.8%**
2P AMD Opteron 1.8GHz 2GB/proc PC2700 4GB Total	**2**	**6.22**	**20617**	**672**	**7.2**	**3.11**	**86.4%**
1P AMD Opteron 1.8GHz 2GB PC2700	**1**	**3.14**	**15400**	**336**	**3.6**	**3.14**	**87.1%**

High-Performance BLAS by Kazushige Goto

- sgemm/dgemm/cgemm/zgemm available today
- Optimized http://www.cs.utexas.edu/users/flame/goto

GOTO results were with 64-bit SuSE 8.1 Linux Professional Edition with NUMA kernel and Myrinet MPIch-gm-1.2.5..10 message passing library.

AMD

BACKUP

AMD64 Processors And Target Systems

AMD

AMD Opteron™ Processor 200 Series:

- 2-way server and workstation proc
- 144-bit DDR interface per CPU: 200 266, 333 MHz*
- Three 16-bit HyperTransport™ links per CPU. Typically, two are used to connect to another CPU and I/O
- 1-MB integrated L2 cache per CPU

Upcoming AMD Athlon™ 64 Processor

- Performance Desktop Processor
- 72-bit DDR interface 200, 266, 333, 400 MHz*
- One 16-bit HyperTranport link

NOTE: The Upcoming AMD Athlon 64 and AMD Opteron are AMD64 processors

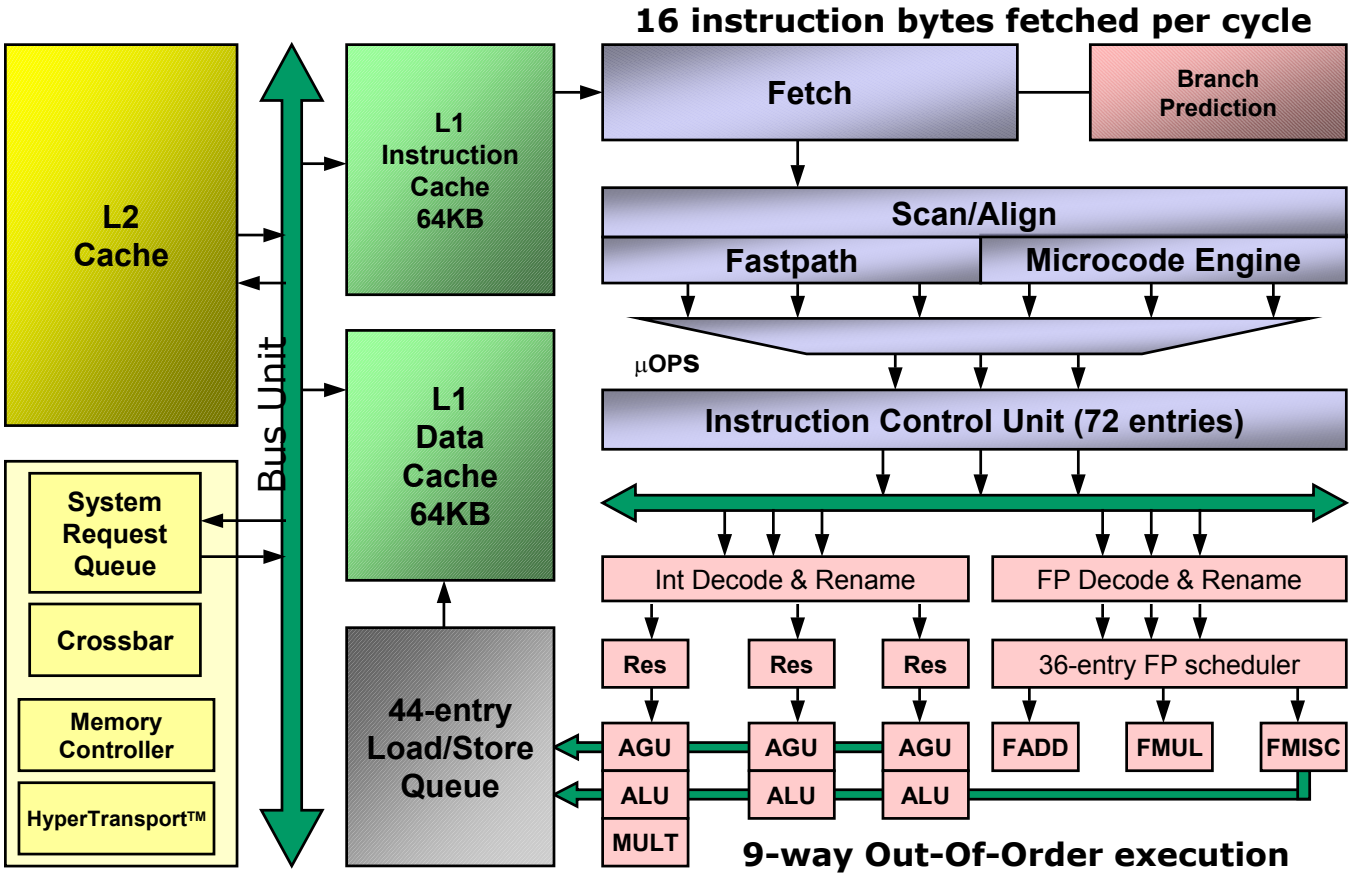
AMD Opteron™ Processor 800 Series:

- Up to 8-way server processor
- 144-bit DDR interface per CPU: 200, 266, 333 MHz*
- Three 16-bit HyperTransport™ Links per CPU. Typically all 3 used to connect to other CPUs & I/O
- 1-MB integrated L2 cache

* = Future memory technology support as it is defined

16-bit HyperTransport Links are at 1600MT/s; provides 6.4GB/s Peak Aggregate Bandwidth

AMD Opteron™ Processor Core Facts



- 12-stage Int, 17-stage fast-path pipelines
- Enhanced TLB structures w/flush filter
- Opts for off-loading writes, probes, memory
- 16 SSE & SSE2 128-bit xmm registers
- 8 legacy x87 80-bit registers

FPU Throughput

- 36 entry FPU instruction scheduler
- 64-bit/80-bit FP Realized thru-put (1 Mul + 1 Add)/cycle: 1.9 FLOPs/cycle
- 32-bit FP Realized thru-put (2 Mul + 2 Add)/cycle: 3.4+ FLOPs/cycle

Compatibility Thunking Layer

- A DLL integral to operating system
  - Transparent to end-user
- Resides within a 32-bit process established by the 64-bit OS to run 32-bit application
- 32-bit application is dynamically linked to Thunking Layer
- Thunking layer implements all 32-bit kernel calls
  - Translates parameters as necessary
  - Calls 64-bit kernel
  - Translates results as necessary
- Well understood technology implemented in Microsoft® Windows®:
  - Windows on Windows (WOW32, WOW64)

Scientific Applications: *Project "Red Storm"*

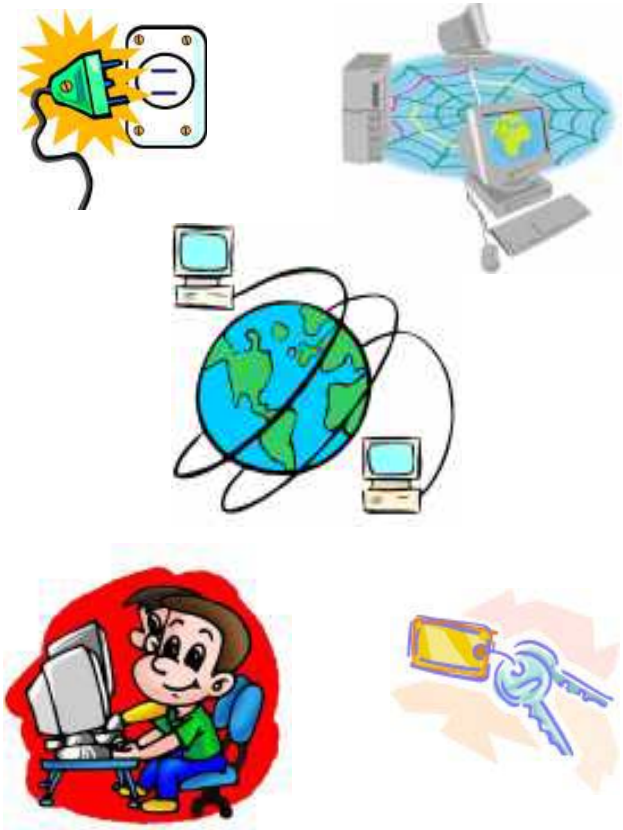
Sandia National Laboratories

CRAY

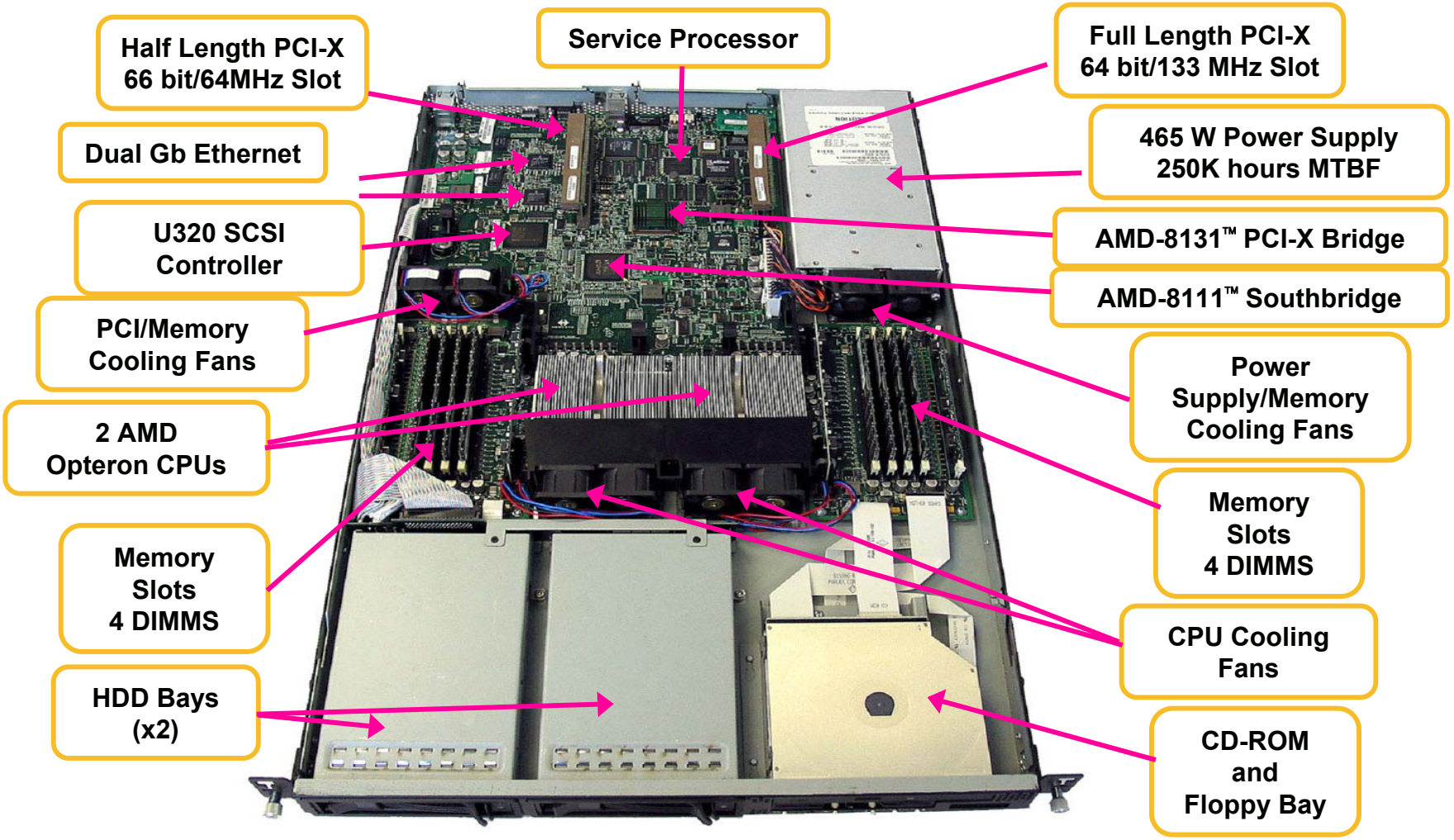
- **Cray Computer plans to build a 40+ teraflop supercomputer using AMD Opteron™ processors for Sandia National Laboratories**
- **Will be used for advanced engineering simulations**
- **$90 million project plans to use more than 10,000 AMD Opteron processors**
- **Will feature a simple building block approach with HyperTransport™ technology that is designed to enable easy implementation and reduce engineering, design, and component costs**

In-Band/Out-of-Band Management

- Eliminates need to physically visit server
  - Remote power down
  - Remote power up (Out-Of-Band only)
  - Hard reset (Out-Of-Band only)
- Available during all server states (setup, boot, OS, or halted)
- Secured against malicious attacks
  - Multi-layer passwords (for remote power features)
  - SSL authentication



Internal View of 1u 2P Server



Other Development Tools

- Absoft will be bringing their full set of FORTRAN toolsets to the AMD64 architecture on both Linux and Windows®
  - Potential beta testers should send email to: opteronbeta@absoft.com
  - Beta available June 2003

- MigraTEC's source code migration tool, 64Express, is now available to aid in the migration of C/C++ code from 32-bit to 64-bit – Available Now
- MigraTEC's cross-platform tool, 32Direct, is now available to assist in cross-platform migrations (i.e. Solaris to Linux) – Available Now

- Etnus has announced 32-bit support of x86-64 with their TotalView distributed debugging product

Other Development Tools

- ATLAS (Automatically Tuned Linear Algebra Subroutines)
- ATLAS has incorporated optimized 64-bit Linux routines to their 3.5.0 Developer release - http://math-atlas.sourceforge.net/ - Available Now
  - Further 64-bit optimizations are forthcoming

- Scyld Computing has announced their intent to support the AMD64 architecture with their Beowulf product around time of AMD Opteron™ processor launch

- MPICH
  - MPICH is available via the open-source community and Linux distributions

Other Development Tools

pallas

- Announced 32-bit support with Vampir/Vampirtrace for the AMD64 architecture

- Announced support for AMD64 with their Distributed Debugger Tool (DDT)
  - This is the first graphical software debugger to support AMD64 with a 64-bit OS
  - Commercial release now available

- Blackdown Java
  - Announced support that their J2SE Version 1.4.2 will support the AMD64 architecture on the Linux OS
    - Blackdown is based on Sun's HotSpot technology

Other Development Tools

- Announced their EnfuZion cluster management product will have support for both 32-bit and 64-bit OSes on the AMD64 architecture

- Announced their support for 64-bit versions of their popular message passing implementations – MPIPro (1.2) and ChaMPIon Pro (2.1)

- Announced the release of the NAGWare F95 compilers for 64-bit Linux
  - Available via www.nag.com/f95AMD

AMD64 Computing Strategy (3)

- BIOS is standard x86 32-bit code
  - Transfer to 64-bit mode is done by OS loader. No extra Requirements.
- Legacy Mode
  - AMD64 processors run any 32-bit legacy OS with leading edge performance
  - Fully compatible with existing 32-bit systems and software
- Compatibility Mode under 64-bit OS
  - 64-bit OS runs existing 32-bit Apps with leading edge performance
  - Processor core provides full x86 compatibility at full speed. No application recompile required, no emulation layer
  - OS provides thunking layer at kernel-call boundary
- 64-bit Mode under 64-bit OS
  - Migrate only where warranted, and at user's pace to fully exploit AMD64
  - Even Apps not needing 64-bit addressing can still enjoy performance enhancements from recompiling into 64-bit

Compatibility Mode

- Provides a mode where existing applications can run unchanged under Long Mode
- Selected on a code-segment basis (CS.L=0)
  - Uses far transfer rather than a full mode switch
    - Faster than mode switch
- Application-level code runs unchanged
  - Legacy segmentation
  - Legacy address and data size defaults
- System aspects use 64-bit mode semantics
  - Interrupts and exceptions use Long Mode handling
  - Paging aspects use Long Mode semantics
    - No support for v86